

FORM PTO-1390
(REV 10-95)

U.S. DEPARTMENT OF COMMERCE AND TRADEMARK OFFICE

ATTORNEY'S DOCKET NUMBER
CFV-005.01TRANSMITTAL LETTER TO THE UNITED STATES
DESIGNATED/ELECTED OFFICE (DO/EO/US)
CONCERNING A FILING UNDER 35 U.S.C. 371

U.S. APPLICATION NO. (If known, see 37 CFR 1.5)

09/171553

INTERNATIONAL APPLICATION NO.
PCT/GB97/01087INTERNATIONAL FILING DATE
18 April 1997 (18.04.97)PRIORITY DATE CLAIMED
19 April 1996 (19.04.96)TITLE OF INVENTION:
PORCINE RETROVIRUSAPPLICANT(S) FOR DO/EO/US
Todd CAMPUS and Douglas HOUSE

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

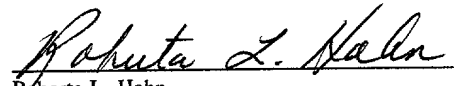
1. (X) This is the FIRST submission of items concerning a filing under 35 U.S.C. 371.
2. () This a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.
3. (X) This express request to begin national examination procedures (35 U.S.C. 371(f) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(l).
4. (X) A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. (X) A copy of the International Application as filed (35 U.S.C. 371(c)(2)).
 - a. () is transmitted herewith (required only if not transmitted by the International Bureau).
 - b. (X) has been transmitted by the International Bureau.
 - c. () is not required, as the application was filed in the United States Receiving Office (RO/US).
6. () A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. () Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)).
 - a. () are transmitted herewith (required only if not transmitted by the International Bureau).
 - b. () has been transmitted by the International Bureau.
 - c. () have not been made; however, the time limit for making such amendments has NOT expired.
 - d. () have not been made and will not be made.
8. () A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. () An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. () A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern document(s) or information included:

11. () An information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. () An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. (X) A FIRST preliminary amendment.
- () A SECOND or SUBSEQUENT preliminary amendment.
14. () A substitute specification.
15. () A change of power of attorney and/or address letter.
16. (X) Other items of information.
 - (X) International Preliminary Examination Report (with translation)
 - (X) PCT Request Form.

Certificate of Express Mail

I hereby certify that the foregoing documents are being deposited with the United States Postal Service as Express Mail, postage prepaid, "Post Office to Addressee", in an envelope addressed to the Assistant Commissioner for Patents, Box PCT, Attn: DO/EO/US, Washington, D.C. 20231 on the date indicated below.



Roberta L. Hahn

Express Mail Label: EL046501414us

Date of Deposit: October 19, 1998

| | | |
|---|--|---|
| U.S. APPLICATION NO. (if known, see 37 CFR 1.5) | INTERNATIONAL APPLICATION NO. PCT/GB97/01087 | ATTORNEY'S DOCKET NUMBER CFV-005.01 |
|---|--|---|

| <p>17. (X) The following fees are submitted:</p> <p>BASIC NATIONAL FEE (37 CFR 1.492(a)(1)-(5)):</p> <p>Search Report has been prepared by the EPO or JPO \$910.00</p> <p>International preliminary examination fee paid to USPTO (37 CFR 1.482) \$700.00</p> <p>No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2)) \$770.00</p> <p>Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO \$1040.00</p> <p>International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4) \$96.00</p> <p style="text-align: right;">ENTER APPROPRIATE BASIC FEE AMOUNT = \$910.00</p> <p>Surcharge of \$130.00 for furnishing the oath or declaration later than () 20 (x) 30 months from the earliest claimed priority date (37 CFR 1.492(e)). \$130.00</p> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 25%;">CLAIMS</th> <th style="width: 25%;">NUMBER FILED</th> <th style="width: 25%;">NUMBER EXTRA</th> <th style="width: 25%;">RATE</th> </tr> <tr> <td>Total claims</td> <td>40 - 20 =</td> <td>20</td> <td>X \$22.00</td> </tr> <tr> <td>Independent claims</td> <td>9 - 3 =</td> <td>6</td> <td>X \$80.00</td> </tr> <tr> <td colspan="3">MULTIPLE DEPENDENT CLAIM(S) (if applicable)</td> <td>+ \$260.00</td> </tr> <tr> <td colspan="3" style="text-align: right;">TOTAL OF ABOVE CALCULATIONS =</td> <td>\$2220.00</td> </tr> <tr> <td colspan="3">Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28).</td> <td>\$0</td> </tr> <tr> <td colspan="3" style="text-align: right;">SUBTOTAL =</td> <td>\$2220.00</td> </tr> <tr> <td colspan="3">Processing fee of \$130.00 for furnishing the English translation later than () 20 () 30 months from the earliest claimed priority date (37 CFR 1.492(f)).</td> <td>\$0</td> </tr> <tr> <td colspan="3" style="text-align: right;">TOTAL NATIONAL FEE =</td> <td>\$2220.00</td> </tr> <tr> <td colspan="3">Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property.</td> <td>\$0</td> </tr> <tr> <td colspan="3" style="text-align: right;">TOTAL FEES ENCLOSED =</td> <td>\$2220.00</td> </tr> </table> <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 70%;"></td> <td style="width: 15%;">Amount to be: refunded</td> <td style="width: 15%;">\$</td> </tr> <tr> <td></td> <td>charged</td> <td>\$</td> </tr> </table> | CLAIMS | NUMBER FILED | NUMBER EXTRA | RATE | Total claims | 40 - 20 = | 20 | X \$22.00 | Independent claims | 9 - 3 = | 6 | X \$80.00 | MULTIPLE DEPENDENT CLAIM(S) (if applicable) | | | + \$260.00 | TOTAL OF ABOVE CALCULATIONS = | | | \$2220.00 | Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28). | | | \$0 | SUBTOTAL = | | | \$2220.00 | Processing fee of \$130.00 for furnishing the English translation later than () 20 () 30 months from the earliest claimed priority date (37 CFR 1.492(f)). | | | \$0 | TOTAL NATIONAL FEE = | | | \$2220.00 | Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property. | | | \$0 | TOTAL FEES ENCLOSED = | | | \$2220.00 | | Amount to be: refunded | \$ | | charged | \$ | <p>CALCULATIONS PTO USE ONLY</p> |
|--|------------------------|--------------|------------------|------|--------------|-----------|----|-----------|--------------------|---------|---|-----------|---|--|--|------------|--------------------------------------|--|--|------------------|---|--|--|-----|-------------------|--|--|------------------|---|--|--|-----|-----------------------------|--|--|------------------|--|--|--|-----|------------------------------|--|--|------------------|--|------------------------|----|--|---------|----|---|
| CLAIMS | NUMBER FILED | NUMBER EXTRA | RATE | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Total claims | 40 - 20 = | 20 | X \$22.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Independent claims | 9 - 3 = | 6 | X \$80.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| MULTIPLE DEPENDENT CLAIM(S) (if applicable) | | | + \$260.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| TOTAL OF ABOVE CALCULATIONS = | | | \$2220.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28). | | | \$0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| SUBTOTAL = | | | \$2220.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Processing fee of \$130.00 for furnishing the English translation later than () 20 () 30 months from the earliest claimed priority date (37 CFR 1.492(f)). | | | \$0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| TOTAL NATIONAL FEE = | | | \$2220.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property. | | | \$0 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| TOTAL FEES ENCLOSED = | | | \$2220.00 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | Amount to be: refunded | \$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | charged | \$ | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

a. (X) A check in the amount of \$2220.00 to cover the above fees is enclosed.

b. () Please charge my Deposit Account No **06-1448**, in the amount of **\$0** to cover the above fees.
A duplicate copy of this sheet is enclosed.

c. (X) The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. **06-1448**, Ref. **CFV-005.01**. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

| | |
|---|--|
| <p>SEND ALL CORRESPONDENCE TO:</p> <p>Patent Group Foley, Hoag & Eliot LLP One Post Office Square Boston, MA 02109-2170</p> | <p>SIGNATURE </p> <p>Chinh H. Pham</p> <p>REGISTRATION NO. 39,329</p> |
|---|--|

09/171553

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

| | | | |
|----------------------------|--------------------------|--------------|------------------|
| Applicant: | Todd CAMPUS et al. | Atty Docket: | CFV-005.01 |
| Serial No.: | Not Yet Assigned | Art Unit: | Not Yet Assigned |
| Date Filed: | Herewith | Examiner: | Not Yet Assigned |
| Based on: | PCT/GB97/01087 | | |
| International Filing Date: | 18 April 1997 (18.04.97) | | |
| Priority Date: | 19 April 1996 (19.04.96) | | |
| Title of Invention: | PORCINE RETROVIRUS | | |

FIRST PRELIMINARY AMENDMENT

Assistant Commissioner for Patents
Washington, D.C. 20231

Sir:

Prior to consideration of the above-referenced application, amend the application as follows:

IN THE CLAIMS:

Please cancel claims 6, 7 and 25 without prejudice.

Please amend the claims as follows:

1. (Amended) An isolated polynucleotide fragment as shown in Figures 1, 2, 3 or 4, subsequence thereof, or corresponding RNA sequence thereof:
 - (a) encoding at least one porcine retrovirus (PoEV) expression product;
 - (b) encoding a derivative of said expression product displaying a physiological [physiologically active] and/or immunological activity [immunogenic derivative] of said expression product as described in (a); or
 - (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).
2. (Amended) An isolated polynucleotide fragment according to claim 1:
 - (a) encoding at least one polypeptide having an amino acid sequence which is shown in Figures 3 or 4 [the polymerase (POL) polypeptide];

- (b) encoding a [physiologically active and/or immunogenic] derivative of said at least one [a] polypeptide displaying a physiological and/or immunological activity substantially similar to the physiological and/or immunological activity of said expression product as described in (a); or
 - (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).
3. (Amended) An isolated polynucleotide fragment according to claim 1 or 2:
- (a) encoding the polymerase (POL) [virion core polypeptide (GAG) and/or envelope] polypeptide [(ENV)];
 - (b) encoding a [physiologically active and/or immunogenic] derivative [of a] polypeptide displaying a physiological and/or immunological activity substantially similar to the physiological and/or immunological activity of the polymerase (POL) as described in (a); or
 - (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).
4. (Amended) An isolated polynucleotide fragment according to claim [1] 2:
- (a) encoding the virion core polypeptide (GAG)[, polymerase (POL) and] and/or envelope polypeptide (ENV);
 - (b) encoding a [physiologically active and/or immunogenic] derivative [of a] polypeptide displaying a physiological and/or immunological activity of said virion core polypeptide (GAG) and/or envelope polypeptide (ENV) as described in (a); or
 - (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

5. (Amended) An isolated polynucleotide fragment [according to any one of claims 1 to 4 wherein the polynucleotide fragments is a deoxyribose nucleic acid (DNA) fragment] displaying at least 90% sequence identity with the sequence as shown in Figures 2 or 3:
- (a) encoding the virion core polypeptide (GAG), polymerase (POL) and envelope polypeptide (ENV) of porcine retrovirus (PoEV);
 - (b) encoding a derivative polypeptide displaying a physiological and/or immunological activity substantially similar to the physiological and/or immunological activity of said GAG POL and ENV polypeptides as described in (a); or
 - (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).
8. (Amended) A recombinant nucleic acid molecule comprising a polynucleotide fragment according to any one of claims 1 to [7] 5.
12. (Amended) A prokaryotic or eukaryotic host cell transformed by a polynucleotide fragment, recombinant nucleic acid molecule, or vector according to any of claims 1 to 5 and 8 to 11.
13. (Amended) A recombinant PoEV polypeptide [or derivative thereof displaying POL PoEV physiological and/or immunogenic] comprising an amino acid sequence as shown in Figure 3 displaying POL activity.
14. (Amended) A recombinant PoEV polypeptide [or derivative thereof displaying GAG and/or ENV PoEV physiological and/or immunogenic activity] comprising an amino acid sequence with at least 95% sequence identity to the GAG amino acid sequence as shown in Figure 3.

15. (Amended) A recombinant PoEV polypeptide [or derivative thereof displaying GAG, POL and ENV PoEV physiological and/or immunogenic activity] comprising an amino acid sequence with at least 75% sequence identity to the ENV amino acid sequence as shown in Figures 3 or 4.
16. (Amended) A recombinant PoEV polypeptide [according to any one of claims 13 to 15] comprising a sequence as shown in Figures 3 or 4, [or functionally active derivative thereof] or derivative polypeptide displaying a physiological and/or immunological activity of the PoEV polypeptide.
18. (Amended) An anti-PoEV antibody or fragment thereof [capable of binding to a polypeptide or fragment] raised against a polypeptide or derivative according to any one of claims 13 or 16.
19. (Amended) A polynucleotide primer which is [PoEV specific] capable of specifically hybridizing to a PoEV polynucleotide fragment as shown in figures 1, 2, 3 or 4 and capable of initiating chain extension from the 3' end of the primer, but which is not able to correctly initiate chain extension from non PoEV sequences.
20. (Amended) A polynucleotide probe which is capable of specifically [hybridising] hybridizing under stringent conditions to a [PoEV] polynucleotide sequence as shown in Figures 1, 2, 3 or 4, but not to non PoEV sequences under stringent conditions.

21. (Amended) A probe or a primer according to claims 19 or 20 which [has] have substantial nucleotide sequence identity with a strand of the molecule depicted in Figures 1, 2, 3 or 4 or a strand complementary therewith, with a corresponding RNA molecule, or with a part of such a molecule.
23. (Amended) Use of a PoEV specific polynucleotide primer or probe according to any of claims 19 to 22 in the detection of PoEV in a sample.
24. (Amended) Use of a PoEV specific polynucleotide [in a PCR] primer or primers according to either of claims 19 or 21 in a polymerase chain reaction for the detection of PoEV in a sample.
26. (Amended) Cells, tissues or organs obtainable from a pig [accoding] according to claim [25] 33.
27. (Amended) [Use of a] A recombinant PoEV polypeptide according to any one of claims 13 to 16 for use in the preparation of a vaccine.
28. (Amended) Use of an anti-PoEV antibody, [a] polynucleotide primer or probe according to any of claims 19 to 21 in the preparation of a detection kit capable of detecting the presence of PoEV nucleic acid in a sample.
29. (Amended) [Use of a] A polynucleotide; polypeptide; antibody; cells, tissues or organs according to any one of claims 1 to [7] 5, 11 to 13, [to] 16 or 26 in therapy or diagnosis.

30. (Amended) Use of a [A] polynucleotide; polypeptide; antibody; cells, tissues or organs according to anyone of claims 1 to [7] 5, 13 to 16 or 26 in the preparation of a medicament for use in therapy or diagnosis.

Please add the following claims:

32. (New) Porcine embryos, embryonic stem cells or cells containing totipotent nuclei capable of forming a viable embryo which have been manipulated by use of a polynucleotide sequence derived from the polynucleotide sequence shown in Figures 1, 2, 3 or 4 so as to substantially prevent or reduce infectious PoEV expression.
33. (New) A pig obtainable from the porcine embryos, embryonic stem cells or cells containing totipotent nuclei capable of forming a viable embryo according to claim 32.

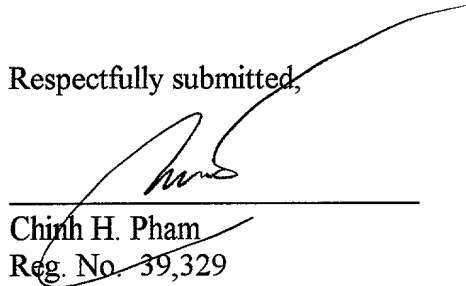
REMARKS

In the present Amendment, Applicants have canceled claims 6, 7 and 25, have added claims 31 and 32, and amended claims 1-5, 8, 12-16, 18-21, 23-24, and 26-30. It is believed that these changes do not introduce any new matter into the claims.

A check in the amount of \$2220 is enclosed to cover the national filing fees, fees for extra claims, and fees for late filing of the Declaration and Oath.

Applicants do not believe that any additional fees are required. However, if any fees, including extensions of time are required, Applicants hereby petition for same and request that the extension or other fee required for timely consideration of this application be charged to Deposit Account No. 06-1448.

Respectfully submitted,



Chinh H. Pham
Reg. No. 39,329

Dated: October 19, 1998

Patent Group
Foley, Hoag & Eliot, LLP
1 Post Office Square
Boston, Massachusetts 02109
Tel.: 617-832-1234
Fax: 617-832-1000

Porcine Retrovirus

The present invention relates inter alia to porcine retrovirus (PoEV) fragments, in particular polynucleotide fragments encoding at least one porcine retrovirus expression product, a recombinant vector comprising at least one polynucleotide fragment, use of PoEV polynucleotide fragments in the detection of native porcine retrovirus, a host cell containing at least one PoEV polynucleotide fragment or a recombinant vector comprising at least one PoEV polynucleotide fragment, PoEV polypeptides, antibodies immuno-reactive with PoEV polypeptides, pharmaceutical compositions comprising recombinant PoEV polypeptides for use as prophylactic and/or therapeutic agents and uses of PoEV polynucleotide fragments and/or polypeptides in medicine, including veterinary medicine and in the preparation of medicaments for use in medicine, including veterinary medicine.

Porcine retrovirus (PoEV) is an endogenous (genetically acquired) retrovirus isolated from pigs and expressed in cell lines derived from porcine material. There are no known pathogenic effects associated with the virus per se in its natural host although the virus appears to be associated with lymphomas in pigs and related viruses are associated with leukaemias and lymphomas in other species. The virus has been reported to infect cells from a variety of non-porcine origins and is, therefore, designated as a xenotropic, amphotropic or polytrophic virus (Lieber MM, Sherr CJ, Benveniste RE and Todaro

GJ. 1975; Strandstrom H, Verjalainen P, Moening V, Hunsmann G, Schwarz H, and Schafer W. 1974; Todaro GJ, Benveniste RE, Lieber MM and Sherr CJ. 1974). The observation that the above viruses may have the potential to infect humans and have a pathogenic effect suggests that the issue of porcine retroviruses must be addressed in the context of xenotransplanting pig tissues or cells. Therefore, information on the properties of PoEV and the development of diagnostic reagents, molecular engineering tools and potential vaccine materials is of paramount importance for example in xenotransplantation technology and the like.

It is an object of the present invention to obviate and/or mitigate against at least some of the above disadvantages.

In one aspect the present invention provides an isolated PoEV polynucleotide fragment:

- (a) encoding at least one porcine retrovirus (PoEV) expression product;
- (b) encoding a physiologically active and/or immunogenic derivative of said expression product; or
- (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

Preferably, the polynucleotide fragment encodes the gag gene (gag), polymerase gene (pol) and/or envelope (env) gene of PoEV. Thus, said expression product can be the virion core polypeptides (GAG) and polymerase (POL) and/or envelope (ENV) polypeptides of PoEV. Thus, the invention further provides a recombinant PoEV virion core, polymerase and/or envelope polypeptide.

"Polynucleotide fragment" as used herein refers to a chain of nucleotides such as deoxyribose nucleic acid (DNA) and

transcription products thereof, such as RNA. Naturally, the skilled addressee will appreciate the whole naturally occurring PoEV genome is not included in the definition of polynucleotide fragment.

The polynucleotide fragment can be isolated in the sense that it is substantially free of biological material with which the whole genome is normally associated *in vivo*. The isolated polynucleotide fragment may be cloned to provide a recombinant molecule comprising the polynucleotide fragment. Thus, "polynucleotide fragment" includes double and single stranded DNA, RNA and polynucleotide sequences derived therefrom, for example, subsequences of said fragment and which are of any desirable length. Where a nucleic acid is single stranded then both a given strand and a sequence complementary thereto is within the scope of the present invention.

In general, the term "expression product" refers to both transcription and translation products of said polynucleotide fragments. When the expression product is a "polypeptide" (i.e. a chain or sequence of amino acids displaying a biological and/or immunological activity substantially similar to the biological and/or immunological activity of PoEV virion core, polymerase and/or envelope protein), it does not refer to a specific length of the product as such. Thus, the skilled addressee will appreciate that "polypeptide" encompasses *inter alia* peptides, polypeptides and proteins of PoEV. The polypeptide if required, can be modified *in vivo* and *in vitro*, for example by glycosylation, amidation, carboxylation, phosphorylation and/or post-translational cleavage.

Polynucleotide fragments comprising portions encompassing the PoEV genome, and derived from retrovirus particles released from a reverse transcriptase-positive porcine kidney cell line PK-15, have been molecularly cloned into a plasmid vector. This was achieved by synthesising cDNAs of PoEV RNA genomes which were recovered from porcine kidney cells expressing the endogenous virus. The cDNA was cloned into a plasmid vector and the isolated PoEV DNA fragment determined (see Figures 1,2 and 3). The sequence of the sequence identified in Figure 1 was the earliest determined sequence, followed by the sequence in Figure 2 and lastly by the most recently revised sequence shown in Figure 3. An additional study has been carried out to determine whether or not the human cell line "Raji" was susceptible to infection with the PoEV present in porcine kidney cells (PK15). A raji clone has now been obtained and the DNA sequence of its env gene region has been determined (see Figure 4).

The DNA fragment of Figure 3 was shown to encode three open reading frames (ORFs) of 524, 1194 and 656 amino acids respectively.

A comparison of the amino acid sequence against previously sequenced retroviruses from other species indicated that novel retrovirus cDNA had been cloned. Sequence analysis using the Lasergene software from DNASTAR Inc. showed that homologies were observed between the cloned PoEV DNA and the majority of retroviruses and that the closest homologies were to gibbon leukaemia virus (GaLV) in the polymerase (pol) and (env) regions of the pro-virus.

The first open reading frame ORF of Figure 3 (nucleotides 588-

2162) is predicted to encode the PoEV virion core polypeptide (*gag* gene). The second ORF (nucleotides 2163-5747) is predicted to encode the PoEV polymerase polypeptide (*pol* gene). The third ORF (nucleotides 5620-7590) is predicted to encode the PoEV envelope polypeptide (*env* gene). The skilled addressee will appreciate that it is possible to genetically manipulate the polynucleotide fragment or derivatives thereof, for example to clone the gene by recombinant DNA techniques generally known in the art and to express the polypeptides encoded thereby in vitro and/or in vivo. DNA fragments having the polynucleotide sequence depicted in Figures 1,2,3 and/or 4 or DNA/RNA derivatives thereof, may be used as a diagnostic tool or as a reagent for detecting PoEV nucleic acid in nucleic acids from donor animals or as a vaccine.

Preferred fragments of this aspect of the invention are polynucleotide fragments encoding: (a) at least one of the one to three polypeptides having an amino acid sequence which is shown in Figures 1,2,3 and/or 4 (b) encoding a polypeptide which is a physiologically active and/or immunogenic derivative of at least one of the polypeptides defined in (a); or (c) which is complementary to a polynucleotide sequence as defined above; or polynucleotide fragments: (a) comprising at least one of the ORFs shown in Figures 1,2,3 and/or 4 or comprising a corresponding RNA sequence; (b) comprising a sequence having substantial nucleotide sequence identity with a sequence as described in (a) above; or (c) comprising a sequence which is complementary to a sequence as described in (a) or (b) above. It is to be understood that the term "substantial sequence identity" is taken to mean at

least 50% (preferably at least 75%, at least 90%, or at least 95%) sequence identity.

The polynucleotide fragment of the present invention may be used to examine the expression and/or presence of the PoEV virus in donor animals and cells, tissues or organs derived from the donor animals to see if they are suitable for xenotransplantation (i.e. PoEV free). In addition, the recipients of pig cells, tissues or organs can be examined for the presence and/or expression of PoEV virus directly or by co-culture or infection of susceptible detector cells.

A polynucleotide fragment of the present invention may be used to identify polynucleotide sequences within the PoEV genome which are PoEV specific (i.e. it is not necessary for the complete PoEV genome to be identified). Such PoEV specific polynucleotide sequences may be used to identify PoEV nucleic acid in samples, such as transplanted cells, tissues or organs and may be included in a definitive test for PoEV.

Thus, the present invention further provides an isolated PoEV polynucleotide fragment capable of specifically hybridising to a PoEV polynucleotide sequence. In this manner, the present invention provides probes and/or primers for use in *ex vivo* and/or *in situ* PoEV virus detection and expression studies. Typical detection studies include polymerase chain reaction (PCR) studies, hybridisation studies, or sequencing studies. In principle any PoEV specific polynucleotide sequence from the above identified PoEV sequence may be used in detection and/or expression studies.

"Capable of specifically hybridising" is taken to mean that

said polynucleotide fragment preferably hybridises to a PoEV polynucleotide sequence in preference to polynucleotide sequences of other virus, animal (especially porcine or human sequences) and/or other species. In a preferred embodiment the PoEV fragment specifically binds to a native PoEV polynucleotide sequence or a part thereof.

The invention includes polynucleotide sequence(s) which are capable of specifically hybridising to a PoEV polynucleotide sequence or to a part thereof without necessarily being completely complementary to said PoEV polynucleotide sequence or fragment thereof. For example, there may be at least 50% preferably at least 75%, most preferably at least 90% or at least 95% complementarity. Of course, in some cases the sequences may be exactly complementary (100% complementary) or nearly so (e.g. there may be less than 10, preferably less than 5 mismatches). Thus, the present invention also provides anti-sense or complementary nucleotide sequence(s) which is/are capable of specifically hybridising to the disclosed DNA sequence. If a PoEV specific polynucleotide is to be used as a primer in PCR and/or sequencing studies, the polynucleotide must be capable of hybridising to PoEV nucleic acid and capable of initiating chain extension from 3' end of the polynucleotide, but not able to correctly initiate chain extension from non PoEV sequences (especially from human, or non-PoEV porcine sequences).

If a PoEV specific test polynucleotide sequence is to be used in hybridisation studies, to test for the presence of PoEV nucleic acid in a sample, the test polynucleotide should preferably remain hybridised to a sample polynucleotide under

stringent conditions. If desired, either the test or sample polynucleotide may be immobilised. Generally the test polynucleotide sequence is at least 10 or at least 50 bases in length. It may be labelled by suitable techniques known in the art. Preferably the test polynucleotide sequence is at least 200 bases in length and may even be several kilobases in length. Thus, either a denatured sample or test sequence can be first bound to a support. Hybridization can be effected at a temperature of between 50 and 70°C in double strength SSC (2xNaCl 17.5g/l and sodium citrate (SC) at 8.8g/l) buffered saline containing 0.1% sodium dodecyl sulphate (SDS). This can be followed by rinsing of the support at the same temperature but with a buffer having a reduced SSC concentration. Depending upon the degree of stringency required, and thus the degree of similarity of the sequences, such reduced concentration buffers are typically single strength SSC containing 0.1%SDS, half strength SSC containing 0.1%SDS and one tenth strength SSC containing 0.1%SDS. Sequences having the greatest degree of similarity are those the hybridisation of which is least affected by washing in buffers of reduced concentration. It is most preferred that the sample and inventive sequences are so similar that the hybridisation between them is substantially unaffected by washing or incubation in one tenth strength sodium citrate buffer containing 0.1%SDS.

PoEV specific oligonucleotides may be designed to specifically hybridise to PoEV nucleic acid. They may be synthesised, by known techniques and used as primers in PCR or sequencing reactions or as probes in hybridisations designed to

detect the presence of PoEV material in a sample. The oligonucleotides may be labelled by suitable labels known in the art, such as, radioactive labels, chemiluminescent labels or fluorescent labels and the like. Thus, the present invention also provides PoEV specific oligonucleotide probes and primers.

The term "oligonucleotide" is not meant to indicate any particular length of sequence and encompasses nucleotides of preferably at least 10b (e.g. 10b to 1kb) in length, more preferably 12b-300b in length and most preferably 15b to 100b.

The PoEV specific oligonucleotides may be determined from the PoEV sequences shown in Figure 1 and may be manufactured according to known techniques. They may have substantial sequence identity (e.g. at least 50%, at least 75%, at least 90% or at least 95% sequence identity) with one of the strands shown therein or an RNA equivalent, or with a part of such a strand. Preferably such a part is at least 10, at least 30, at least 50 or at least 200 bases long. It may be an ORF or a part thereof.

Oligonucleotides which are generally greater than 30 bases in length should preferably remain hybridised to a sample polynucleotide under one or more of the stringent conditions mentioned above. Oligonucleotides which are generally less than 30 bases in length should also preferably remain hybridised to a sample polynucleotide but under different conditions of high stringency. Typically the melting temperature of an oligonucleotide less than 30 bases may be calculated according to the formula of; 2°C for every A or T, plus 4°C for every G or C, minus 5°C . Hybridisation may take place at or around the calculated melting temperature for any particular

oligonucleotide, in 6 x SSC and 1% SDS. Non specifically hybridised oligonucleotides may then be removed by stringent washing, for example in 3 x SSC and 0.1% SDS at the same temperature. Only substantially similar matched sequences remain hybridised i.e. said oligonucleotide and corresponding PoEV nucleic acid.

When oligonucleotides of generally less than 30 bases in length are used in sequencing and/or PCR studies, the melting temperature may be calculated in the same manner as described above. The oligonucleotide may then be allowed to anneal or hybridise at a temperature around the oligonucleotides calculated melting temperature. In the case of PCR studies the annealing temperature should be around the lower of the calculated melting temperatures for the two priming oligonucleotides. It is to be appreciated that the conditions and melting temperature calculations are provided by way of example only and are not intended to be limiting. It is possible through the experience of the experimenter to vary the conditions of hybridisation and thus anneal/hybridise oligonucleotides at temperatures above their calculated melting temperature. Indeed this can be desirable in preventing so-called non-specific hybridisation from occurring.

It is possible when conducting PCR studies to predict an expected size or sizes of PCR product(s) obtainable using an appropriate combination of two or more PoEV oligonucleotides, based on where they would hybridise to the sequence in Figure 1. If, on conducting such a PCR on a sample of PoEV DNA, a fragment of the predicted size is obtained, then this is predictive that

and genomic loci (e.g. non-expressed provirus loci). Such a screening method would facilitate, for example, screening in a population of animals which are bred to lack expressed provirus and genomic loci and/or loci that do not encode infectious virus particles.

Reagents may also be developed from said polynucleotide fragments as aids to develop pigs that do not express an infectious, PoEV capable of infecting humans. Such pigs could still contain partial defective genomes that could result in the expression of non-infectious particles, viral proteins or viral mRNA. Alternatively, it may be possible to use constructs derived from the PoEV polynucleotide sequence to act as insertional mutagens to knockout the productive infectious PoEV in embryos, embryonic stem cells, or cells containing totipotential nuclei capable of forming a viable embryo. Thus gag, pol and/or env gene "knockouts" may be constructed to allow development of breeding programmes in pigs whereby endogenous PoEV is substantially prevented or reduced. For example the nucleotide sequence of PoEV can be manipulated e.g. by deletion of a coding sequence in vitro and the resulting construct used to replace the natural PoEV sequence by recombination. Thus, the proviral genome can be rendered inactive in the porcine cells. The knockouts can be manipulated into embryos and/or stem cells and if required manipulated nuclei can be transferred from target cells to germ cells using micromanipulation techniques well known in the art. The invention also extends to animals derived from such germ cells.

Thus, transgenic pigs may be produced containing anti-sense

constructs and/or ribozyme constructs capable of downregulating the expression of viral proteins, or transgenic pigs expressing a single chain immunoglobulin molecule with specificity for PoEV proteins or other protein that might interfere with protein synthesis or viral assembly may also be produced. Similar transgenes encoding trans-dominant negative regulators of PoEV expression or transgenes encoding competitive defective "genomic RNAs" may be used to reduce or eliminate the production of infectious virions. The generation of reagents to suppress the expression of native PoEV loci in pigs, such as, by generation of antisense nucleic acids (e.g. antisense mRNAs), ribozymes or other antiviral reagents may also be developed.

The polynucleotide fragment can be molecularly cloned into a prokaryotic or eukaryotic expression vector using standard techniques and administered to a host. The expression vector is taken up by cells and the polynucleotide fragment of interest expressed, producing protein. Presentation of the protein on cell surface stimulates the host immune system to produce antibodies immunoreactive with said protein as part of a defence mechanism. Thus, expressed protein may be used as a vaccine.

Inactivated vaccines can be produced from PoEV's or cells releasing PoEV. Such infected cells may be generated by natural infection or by transfection of a proviral clone of PoEV. It will be understood that a proviral clone is a molecular clone encoding on at least one antigenic polypeptide of PoEV. After harvesting the virus and/or the infected cells, viruses or infected cells present can be inactivated for example, with formaldehyde, gluteraldehyde, acetylenimine or other

suitable agent or process to generate an inactivated vaccine using methods commonly employed in the art. (CVMP Working Party on Immunological Veterinary Medicinal Products (1993). General requirements for the production and control of inactivated mammalian bacterial and viral vaccines for veterinary use). Sub unit vaccines may be prepared from the individual proteins encoded by the gag, pol and env genes. Typically a vaccine would contain env gene products either alone or in combination with gag genes produced by expression in bacteria, yeast or mammalian cell systems.

Proviral clones of PoEV can be engineered to develop single cycle or replication defective viral vectors suitable for vaccination using techniques. Such viral vectors known in the art (e.g. MuLV Murine Leukaemia Retrovirus, Adenovirus and Herpesviruses (Anderson WF. (1992). Human Gene Therapy. *Science* 256, 808-813) may have one or more genes essential for replication deleted, with the missing gene function expressed constitutively or conditionally from a further, different construct which is integrated into the chromosomal DNA of a complementing cell line to the proviral PoEV clone. PoEV virions released from the cell line may infect secondary target cells in the vaccinee but not produce further infectious virus particles. For instance, the polynucleotide sequence encoding the reverse transcriptase domain of pol can be deleted from the proviral PoEV clone and the reverse transcriptase domain of pol integrated into the complementing cell line.

It will be understood that the polynucleotides; polypeptides; PoEV free cells, tissues and/or organs encompassed

by the present invention could be used in therapy, diagnosis, and/or methods of treatment. The polynucleotides; polypeptides; PoEV free cells, tissues and/or organs encompassed by the present invention can also be used in the preparation of medicaments for use in therapy or diagnosis.

The cloning and expression of a recombinant PoEV polynucleotide fragment also facilitates in producing anti-PoEV antibodies and fragments thereof (particularly monoclonal antibodies) and evaluation of in vitro and in vivo biological activity of recombinant PoEV polymerase and/or envelope polypeptides. The antibodies may be employed in diagnostic tests for native PoEV virus.

It will be understood that for the particular PoEV polypeptides embraced herein, natural variations can exist between individuals or between members of the family Suidae (i.e. the pig family). These variations may be demonstrated by (an) amino acid difference(s) in the overall sequence or by deletions, substitutions, insertions, inversions or additions of (an) amino acid(s) in said sequence. All such derivatives showing active polymerase and/or envelope polypeptide physiological and/or immunological activity are included within the scope of the invention. For example, for the purpose of the present invention conservative replacements may be made between amino acids within the following groups:

- (I) Alanine, serine, threonine;
- (II) Glutamic acid and aspartic acid;
- (III) Arginine and leucine;
- (IV) Asparagine and glutamine;

(V) Isoleucine, leucine and valine;

(VI) Phenylalanine, tyrosine and tryptophan

Moreover, recombinant DNA technology may be used to prepare nucleic acid sequences encoding the various derivatives outlined above.

As is well known in the art, the degeneracy of the genetic code permits substitution of bases in a codon resulting in a different codon which is still capable of coding for the same amino acid, e.g. the codon for amino acid glutamic acid is both GAT and GAA. Consequently, it is clear that for the expression of polypeptides with the amino acid sequences shown in Figure 1 or fragments thereof, use can be made of a derivative nucleic acid sequence with such an alternative codon composition different from the nucleic acid sequence shown in said Figure 1.

Furthermore, fragments derived from the PoEV core, polymerase and/or envelope polypeptides as depicted in Figure 3, which still display PoEV virus core polypeptide, polymerase and/or envelope polypeptide properties, or fragments derived from the nucleic acid sequence encoding the virus core polypeptides, polymerase and/or envelope polypeptides or derived from the nucleotide sequence depicted in Figures 1,2,3 and/or 4 encoding fragments of said virus core polypeptide, polymerase and/or envelope polypeptides are also included of the present invention. Naturally, the skilled addressee will appreciate within the ambit that the said fragments should substantially retain the physiological and/or immunological properties of the GAG, POL and/or ENV polypeptides.

The PoEV polynucleotide fragment of the present invention

is preferably linked to regulatory control sequences. Such control sequences may comprise promoters, operators, inducers, enhancers, ribosome binding sites, terminators etc. Suitable control sequences for a given host may be selected by those of ordinary skill in the art.

A polynucleotide fragment according to the present invention can be ligated to various expression controlling sequences, resulting in a so called recombinant nucleic acid molecule. Thus, the present invention also includes an expression vector containing an expressible PoEV nucleic acid molecule. The recombinant PoEV nucleic acid molecule can then be used for the transformation of a suitable host. Such hybrid molecules are preferably derived from, for example, plasmids or from nucleic acid sequences present in bacteriophages or viruses and are termed vector molecules.

Specific vectors which can be used to clone nucleic acid sequences according to the invention are known in the art (e.g. Rodriguez, R.L. and Denhardt, D.T., Edit., Vectors: a survey of molecular cloning vectors and their uses, Butterworths, 1988).

The methods to be used for the construction of a recombinant nucleic acid molecule according to the invention are known to those of ordinary skill in the art and are *inter alia* set forth in Sambrook, et al. (Molecular Cloning: a laboratory manual Cold Spring Harbour Laboratory, 1989).

The present invention also relates to a transformed cell containing the PoEV polynucleotide fragment in an expressible form. "Transformation", as used herein, refers to the introduction of a heterologous polynucleotide fragment into a

host cell. The method used may be any known in the art, for example, direct uptake, transfection transduction or electroporation (Current Protocols in Molecular Biology, 1995. John Wiley and Sons Inc). The heterologous polynucleotide fragment may be maintained through autonomous replication or alternatively, may be integrated into the host genome. The recombinant nucleic acid molecules preferably are provided with appropriate control sequences compatible with the designated host which can regulate the expression of the inserted polynucleotide fragment, e.g. tetracycline responsive promoter, thymidine kinase promoter, SV-40 promoter and the like.

Suitable hosts for the expression of recombinant nucleic acid molecules may be prokaryotic or eukaryotic in origin. Hosts suitable for the expression of recombinant nucleic acid molecules may be selected from bacteria, yeast, insect cells and mammalian cells.

Since the biological half life and the degree of glycosylation of recombinant PoEV virus core polypeptide, polymerase and/or envelope polypeptides may be important for use *in vivo*, yeast and baculovirus systems, in which a greater degree of processing and glycosylation occur, are preferred. The yeast strain *Pichia Pastoris* exhibits potential for high level expression of recombinant proteins (Clare et al., 1991). The baculovirus system has been used successfully in the production of type 1 interferons (Smith et al., 1983).

Embodiments of aspects of the present invention will now be described by way of example only which are not intended to be limiting thereof.

Examples Section

Example 1

Preparation of viral RNA

500ml of supernatant derived from exponentially growing porcine kidney cells (PK-15, American Type Culture Collection CCL 33) was clarified by centrifugation of approximately 11,000xg for 10 minutes. Virus was pelleted from the clarified supernatant by centrifugation at approximately 100,000xg for 60 minutes. The supernatant was discarded and the viral pellet retained for the preparation of viral RNA genomes. RNA was prepared from the virus pellet using a Dynabeads (registered trade mark) mRNA Direct kit according to the manufacturer's protocols; A PoEV virus pellet was resuspended in 500µl of TNE (10mM Tris HCl pH8.0, 0.1M NaCl, 1mM EDTA) and the virions disrupted by the addition of 2ml of lysis/binding buffer. Dynabeads Oligo(dT)₂₅ were conditioned according to the manufacturer's instructions and added to the virus disrupted solution. Viral RNA was allowed to bind to the Dynabead for 10 minutes before the supernatant was removed and the bound RNA was washed three times with washing buffer with LiDS (0.5ml) and twice with washing buffer alone. The RNA was finally resuspended in 25 µl of elution solution. All procedures were performed at ambient temperature. RNase contamination was avoided by the wearing of gloves, observation of sterile technique and treatment of solutions and non-disposable glass and plasticware with diethyl pyrocarbonate (DEPC). The RNA was resuspended in DEPC- treated sterile water.

Example 2

Synthesis of cDNA

cDNA was synthesised from the purified genomic RNA using Great Lengths TM cDNA amplification reverse transcriptase reagents (Clontech Laboratories Inc.) following the manufacturer's instructions. The RNA was primed with both oligo(dT) and random hexamers to maximise synthesis.

The Great Lengths cDNA synthesis protocol is based on a modified Gubler and Hoffman (1983) protocol for generating complementary DNA libraries and essentially consists of first-strand synthesis, second strand synthesis, adaptor ligation, and size fractionation.

First strand synthesis: lock-docking primers anneal to the beginning of the poly-A tail of the RNA due to the presence of A, C or a residue at the 3'-end of the primer. This increases the efficiency of cDNA synthesis of eliminating unnecessary reverse transcription of long stretches of poly-A. In addition, the reverse transcriptase used is MMLV (RNase H⁻) which gives consistently better yields than do wild-type MMLV or AMV reverse transcriptase.

Second strand synthesis: the ratio of DNA polymerase I for RNase H has been optimised to increase the efficiency of the second strand synthesis and to minimize priming by hair pin loop formation. Following second-strand synthesis, the ds cDNA is treated with T4 DNA polymerase to create blunt ends.

Adaptor ligation: the cDNA is ligated to a specially designed adaptor that has a pre-existing EcoRI "sticky end". The use of this adaptor, instead of a linker, eliminates the need to methylate and the EcoRI - digest the cDNA, and thus leaves internal EcoRI, sites intact. The adaptor is 5'-phosphorylated at the blunt end to allow efficient ligation to the blunt-ended cDNA.

Size fractionation: the ds cDNA is phosphorylated at the EcoRI sites and size-fractionated to remove unligated adaptors and unincorporated nucleotides. The resulting cDNA is ready for cloning into a suitable EcoRI-digested vector.

Example 3

Molecular cloning of cDNA

The size fractionated fragment was ligated with EcoR I- digested pZERO™ -1 plasmid vector DNA (Invitrogen Corporation, San Diego, U.S.). The ligation mix was used to transform competent TOP10F' cells and these were plated onto L-Agar containing zeocin following the manufacturer's instructions (Zero Background™ cloning kit - Invitrogen). Several of the resulting zeocin resistant colonies were amplified in L-Broth containing zeocin and the plasmid DNA was purified by alkaline lysis (Maniatis et al., 1982).

The plasmid DNA was digested to completion with the endonuclease EcoR I and the resulting DNA fragments were separated by electrophoresis through an 1.0% agarose gel (Maniatis et al., 1982), in order to check that a fragment in the

predicted size fractionated size range had been cloned. A clone identified as pPoEV was used in further experimentation.

Example 4

DNA sequence analysis.

pPoEV plasmid DNA was purified according to common techniques (Sambrook et al, 1989) and sequenced using an ABI automated sequencer. Overlapping sequencing primers from both strands of the molecular clone were used to determine the nucleotide sequence.

The first sequence obtained is shown in Figure 1. This sequence was identified as encoding two ORFs of 924 (nucleotides 23-2793) and 218 (nucleotides 2642-3297) amino acids, relating to the *pol* and *env* genes respectively. This sequence was revised and updated to the second sequence as shown in Figure 2. This second sequence was identified as encoding three ORFs of 516 (nucleotides 576-2126), 1186 (nucleotides 2143-5733) and 656 (nucleotides 5606-7576) amino acids, encoding the PoEV *gag*, *pol* and *env* genes respectively. This second sequence has since been revised and updated to the sequence shown in Figure 3. This third sequence was identified as encoding three ORFs of 524 (nucleotides 588-2162), 1194 (nucleotides 2163-5747) and 656 (nucleotides 5620-7590) amino acids, encoding the PoEV *gag*, *pol* and *env* genes respectively.

The differences in the disclosed sequences is reflected by improvements in carrying out and analysing the sequence obtained. However, there is 100% identity at the nucleic acid level, between positions 21-2681 of the first sequence and positions 2972-5653 of the third sequence. Overall there is a 70.5%

identity in the entire 3310 bp of the first sequence with a corresponding portion of the third sequence.

There are only 3 base changes between the second sequence and the third sequence. These are as follows:

| <u>base no. (from Figure 2)</u> | <u>change</u> |
|---------------------------------|--------------------|
| 2121 | insertion of a "G" |
| 2157 | insertion of a "G" |
| 5902 | "R" to an "A" |
| 7700 | "M" to an "A" |

The changes at base nos. 5902 and 7700 do not effect the corresponding amino acid sequence. However, the changes at positions 2121 and 2157 alter the amino acid sequence at the end of GAG and the begining of POL. For GAG the final amino acid "S" have now been replaced by "VLAEEDKD". The total product size is now 524 amino acids. For POL, the first five amino acids "RLGET" have been delated and replaced by "GRR". The total product size is now 1194 amino acids.

Similarities were observed between pPoEV and the majority of retroviruses determined by using alogrithims from DNASTAR Inc. Lasergene software (DNASTAR). The similarities were closest with gibbon ape leukaemia virus (GaLV) in the polymerase (*pol*) regions of the pro-virus at 68.5%, in the virus core (*gag*) region, 59.2% and in the envelope (*env*) region, 39.3% The nucleotide sequence and major ORFs of the pPoEV insert are shown in Figure 3. The largest ORF (nucleotides 2163-5747) encodes the polymerase polypeptide and the smaller ORFs (nucleotides 588-2162 and 5620-7590) encode the core and envelope polypeptides respectively.

Example 5

Purification of cellular DNA from cultured cells, tissues and blood.

Cultured cells

Cells were maintained in culture and approximately 5×10^7 cells were harvested for DNA preparation. The cells were pelleted by centrifugation resuspended in phosphate-buffered saline, re-centrifuged at 1000g for 2 minutes and the supernatant was discarded.

Porcine tissues

Porcine tissue samples were frozen in liquid nitrogen and powdered by grinding in a mortar or between metal foil. The samples were resuspended in 5ml of extraction buffer consisting of 0.025M EDTA (pH 8.0), 0.01M Tris.Cl pH 8.0, 0.5% SDS 20 μ g/ml RNase and 100 μ g/ml proteinase K (Maniatis et al., 1982).

Porcine blood

A buffy coat was prepared from the blood samples. 20ml samples were centrifuged at 1000g for 15 minutes. The buffy coat was resuspended in buffer and the samples centrifuged at 1000g for 15 minutes. The process was repeated one further time. The sample was mixed with 5ml (3x volume) of extraction buffer (Maniatis et al., 1982).

Purification

The samples (i.e. cultured cells, porcine tissue or porcine blood cells) in proteinase K-extraction buffer containing 20 μ g/ml RNase

and 100µg/ml proteinase K were digested for approximately 24 hours at 37°C. The deproteinised DNA was extracted twice with phenol and twice with phenol chloroform and finally precipitated by ethanol in the presence of ammonium acetate. The DNA was recovered by centrifugation at 3000g for 30 minutes and the supernatant discarded (Maniatis et al., 1982). The pellet was washed in 70% ethanol and allowed to air dry for approximately 1 hour. The DNA was allowed to re-dissolve in Tris EDTA (TE) buffer and the purity and concentration of the DNA was assessed by spectrophotometry (Maniatis et al., 1982).

Example 6

Southern blot analysis of porcine tissue and cells

In order to demonstrate that the molecularly cloned DNA comprising the insert from PoEV was derived from the PK-15 cell line (American Type Culture Collection CCL133), the DNA was hybridised against cellular DNAs and its ability to detect proviral DNA was examined.

DNA purified from pPoEV was radioactively labelled and used to probe a Southern blot of endonuclease digested DNAs derived from PK-15 cells .

The DNAs probed were as follows :

- a) Copy number controls of pPoEV DNA linearized by digestion with EcoRI. One copy per haploid cell genome was estimated to be 6.84pg. The control was present at an estimated copy number of 1, 5 and 10 copies.
- b) PK-15 DNA.
- c) Negative control HeLa (American Type Culture Collection

- CCL2) DNA derived from a human adenocarcinoma cell line harbouring human papillomavirus type 18 DNA.
- d) Negative control SP20 (European Collection of Animal Cell Cultures 85072401) DNA derived from a murine myeloma cell line harbouring a xenotropic MuLV retrovirus.

A hybridisation signal was observed in only the PK-15 porcine DNA. No signal was detected in either the negative human or murine DNAs. The PK-15 DNA contained more than 10 copies per cell with an estimated copy number of 20. The sizes of the three major EcoRI- endonuclease digested DNA fragments were approximately 3.8kb, 1.8kb and 0.6kb. The sizes of relevant fragments detected in the recombinant pPoEV were comparable.

There are, as expected, a number of fragments common to the genomic DNA of PK-15 and pPoEV DNA and there is agreement between the patterns observed in both DNAs digested with XhoI, BamHI and HindIII. However, there are additional fragments obtained on digestion of pPoEV DNA by a number of endonucleases.

pPoEV sequences were also detected in swine testes (American Type Culture Collection CRL 1746) and primary porcine kidney cells (Central Veterinary Laboratory batch C04495) but not in hamster CHOK1 (American Type Culture Collection CCL61) or murine NS0 myeloma cells (European Collection of Animal Cell Cultures 85110503).

In order to demonstrate that the molecularly cloned DNA comprising the insert from pPoEV could detect sequences in porcine cells and tissues in addition to PK-15 the pPoEV DNA was hybridised against cellular DNA from tissues derived from pigs and its ability to detect proviral DNA was examined (Maniatis et al., 1982).

The DNA purified from pPoEV was radioactively labelled and used to probe a Southern blot of endonuclease digested DNAs derived from pig organs including liver, kidney, heart and blood.

The DNAs probed were as follows :

- a) Copy number controls of pPoEV DNA linearized by digestion with EcoRI. One copy per haploid cell genome was estimated to be 6.84pg. The control was present at an estimated copy number of 5, 10, 20 and 50 copies.
- b) DNA purified from the porcine tissues digested with EcoRI.

A hybridisation signal was observed in all the porcine DNAs.

The DNAs contained less than 5 copies per cell. There were approximately eight distinct bands in each DNA. The sizes of the three major endonuclease digested DNA fragments were approximately 3.8kb, 1.8kb and 0.6kb.

The PCR reaction amplified a sequence of approximately 787bp from pPoEV and from porcine cells as expected indicating that the assay detected the PoEV proviral DNA. There was no specific amplification of the expected sequence in cells of non-porcine origin and therefore, the PCR reaction and recombinant clone can be used as a specific and sensitive diagnostic tool for detection of PoEV.

Two further digonucleotides were designed against the 3' end of the pol gene and 5' end of the gag gene respectively.

The 3' pol oligonucleotide was 5'-GAT GGC TCT CCT GCC CTT TG-3'

The 5' gag oligonucleotide was 5'-CGA TGG AGG CGA AGC TTA AGG-3'

The above oligonucleotides were also used in PCR reactions according to the conditions described above, with the exceptions that the annealing temperature was 58° and 30 cycles of replication were carried out. The PCR reaction amplified a sequence of approximately 468bp from pPoEV and from porcine cells.

Example 8

Production of PoEV polypeptide in *Escherichia coli*.

The open reading frame (ORF) encoding the pol peptide was isolated from the pPoEV clone and molecularly cloned into the plasmid pGEX-4T-1 (Pharmacia Ltd.) for expression.

Two ml cultures of *E. coli* transformed with various expression constructs were grown with shaking at 37°C to late log phase

(O.D._{600nm} of 0.6) and induced by the addition of IPTG to 0.1 mM. Induced cultures were then incubated for a further 2 hours after which the bacteria were collected by centrifugation. The bacterial pellet was lysed by boiling in SDS-PAGE sample buffer and the protein profile of the induced bacteria was analysed on a 12% acrylamide gel (Laemmli, 1970) followed by staining with coomassie brilliant blue dye.

Example 9

Isolation and partial sequencing of Raji clone

The aim of the study was to determine whether the human cell line "Raji" was susceptible to infection with the PoEV present in porcine kidney cells (PK15). In order to test the capacity of the virus for xenotropism, PK15 cells were co-cultured with the B lymphoblastoid (Raji) cell line over 20 passages.

The culture system utilised direct culture and transwells, which separated the human and porcine cells, but permitted viruses to pass through the separating membrane. After every fifth passage, supernatants from the human cell lines are tested for the presence of retrovirus by reverse transcriptase assay.

Cell cultures

Porcine kidney (PK15) cells (ATCC CCL 33) were used as the source of PoEV. The human cells used for co-cultivation with PK15 cells were the lymphoblast-like Burkitts lymphoma Raji (ATCC CCL 86) cell line. This cell line does not harbour endogenous

retroviruses and lacks reverse transcriptase activity when tested by the present inventors.

Co-cultivation

Raji cells were co-cultivated directly with PK15 cells in duplicate 80cm² flasks and exposed to the PK15 cells throughout the 20 passage culture period. The cells were passaged twice per week and PK15 cells added as necessary from a stock culture. At every fifth passage a sample of Raji cells was removed from the co-culture, washed and cultured for 3-4 days. Supernatant was then harvested and tested for presence of retrovirus by reverse transcriptase (RT) assay.

RESULTS

The presence of reverse transcriptase activity with a preference for the Mn²⁺ cation in the supernatant from detector cell cultures is suggestive of infection by porcine retrovirus. Reverse transcriptase activity with preference for the Mn²⁺ template was not detected in the duplicate co-cultivated test cultures at passage 5 but was detected at passages 10, 15 and 20. No significant RT activity was detected in the negative control cultures. RT activity with preference for the Mn²⁺ template was detected in positive control cultures at passage 5 and 20.

An infected raji culture was diluted to single cells, and then a selection of cells cultured separately such that each culture originated from one cell. Each culture was tested by reverse-transcriptase assay.

Genomic DNA was made from an RT-positive clone as described in example 5 -purification. The PoEV ENV region was amplified by PCR as described below and the product molecularly cloned into pMOS blue T-vector (Amersham). This molecular clone was then sequenced (Fig. 4).

PCR

Oligonucleotides were selected from the PoEV genome.

The upstream primer was 5'-GAT GGC TCT CCT GCC CTT TG -3'

5' base position: 5240

The downstream primer was 5'-CCA CAG TCG TAC ACC ACG -3'

5' base position: 8144

Expected product size: 2904bp

Approx. 1 μ g of genomic raji clone DNA was added to a 50 μ l reaction mixture containing 200 μ M of dATP, dCTP, dGTP, dTTP, 30pM each primer detailed above, 1u Taq DNA polymerase and 5 μ l reaction buffer. The reaction buffer contained 200mM Tris.Cl pH 8.4, 500mM potassium chloride, 15mM magnesium chloride and ultrapure water. The solution was overlaid with two drops of mineral oil to prevent evaporation. Thirty cycles of amplification was performed followed by an elongated extension reaction of 60min. at 72°C.

The cycles consisted of:

95°C 1 min.

56°C 1 min.

72°C 2 min.

The PCR product was visualised as described in example 7.

Product size: ~3Kb.

CLONING

The PCR product was molecularly cloned into pMOS-Blue T-vector as directed by the manufacturer (pMOS-Blue T-vector kit - Amersham).

20 transformed colonies (clones) were picked and added to 5mls L-broth containing 50 µg/ml ampicillin. The cultures were grown shaking at 37°C overnight. Plasmid DNA was isolated from each clone using the perfect prep plasmid isolation kit as directed by the manufacturer (5 Prime - 3 Prime Inc. Boulder, CO, USA).

Plasmid DNA was digested to completion with the endonucleases EcoRI and HindIII and the products visualised on an ethidium bromide-stained 1% agarose gel. A clone (raji env clone) showing the same banding pattern as that predicted for 'PK15 cell line derived PoEV', was selected for sequencing.

SEQUENCING

Raji env clone plasmid DNA prepared above was sequenced using an ABI automated sequencer, and the commercially available T7 sequencing primer.

The entire env gene region of the "Raji" was sequenced (see Figure 4) and discovered to have substantial sequence identity at both the nucleic acid and amino acid levels (98.9% and 96.3% respectively) with the PoEV sequence from PK-15.

Example 10

Phylogenetic analysis

Phylogenetic analysis was performed using the PHYLIP package. Sequence distances were calculated using the PROTDIST program (Dayhoff matrix) and a neighbour-joining unrooted phylogenetic tree reconstructed using the NEIGHBOUR program.

Bootstrapping was performed using 200 replicates of the pol alignment, created using the SEQBOOT program and a consensus tree was obtained using the CONSENSE program (see Figure J). The bootstrap percentages are indicated at the branch fork, with missing values equal to 100%. The data indicate that PoEV is closely related to but distinct from the type-C oncovirus typified by gibbon, murine and feline leukaemia viruses.

A phylogenetic tree was constructed from the pol alignment using the maximum likelihood algorithm (Dayhoff matrix). This tree differed from the pol NJ tree only in the placement of the BaEV lineage in relation to other mammalian type C viruses and corresponded to a low bootstrap for the BaEV fork observed in the NJ tree.

Example 11Analysis of the LTR and adjacent region

The long terminal repeat (LTR) is a reiterated sequence at each end of the provirus that contains the enhancer and promoter governing transcription of the provirus as well as sequences required for reverse transcription of the RNA genome and integration of the proviral DNA. Three recognised domains of the LTR are identifiable, U3, R and U5 with the LTR being delineated by inverse repeats AATGAAAGG and CCTTTCATT at the 5' and 3' ends of U3 and U5 respectively.

| <u>LTR Domain</u> | <u>PoEV Genome Sequence</u> | <u>Length bp</u> |
|-----------------------------|-----------------------------|------------------|
| in accordance with Figure 3 | | |
| U3 | 7638-8106 | 469 |
| R* | 8107-8138, 1-61 | 82 |
| U5 | 62-143 | 82 |

*The position of the R is defined here by similarity to the 3' end of the MuLV LTR and is compatible with the observed location of a cap site approximately 24 bp downstream of the TATA box.

The U3 region contains multiple potential transcription sites as shown in Figure 6. Most of the U3 region shows little or no homology to other mammalian type-C retroviruses which show conserved sites or repeat elements. However, there is homology to other mammalian type-C viruses towards the 3' end of the U3 & region and into R and U5. Amongst the potential transcription factor sites are those for the following:

LyF-1 is a transcriptional regulator that interacts with a novel class of promoters for lymphocyte-specific-genes (Lo et al 1991).

E47 is the prototype member of a new family of tissue specific enhancer proteins that have been shown to bind to the enhancer of murine leukaemia virus.

ETS-1 is a transcription factor primarily expressed in the haematopoietic lineage.

The LTR contains direct repeats at 80006-8062 and 8045-8101 which together contain three potential CCAATT boxes. A potential TATA box is located at position 8129-8144.

The R region contains a PADS (Poly A downstream element) and consensus polyadenylation signal (AATAAA).

The primer binding site (PBS) of PoEV is glycine(2) tRNA which has not reported for any exogenous retrovirus.

References

- Clare JJ, Rayment FB, Ballantine SP, Sreekrishna K and Romanos MA. (1991). High Level expression of tetanus toxin fragment C in *Pichis pastoris* strains containing multiple tandem integrations of the gene. *Biototechnology*, 9, 455-460
- Derynck R, Singh A and Goeddel DV. (1983). Expression of the human interferon- γ in yeast. *Nucleic Acids Res.*, 11, 1819-1837.
- DNASTAR. (1994). Lasergene Biocomputing Software for Windows. User's Guide.
- Invitrogen. Version A. Zero Background™ Cloning Kit Catalog no K2500-01.
- Laemmli UK. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4N. *Nature*, 227, 680-685.
- Lieber MM, Sherr CJ, Benveniste RE and Torado GJ. (1975). Biologic and immunologic properties of porcine type C viruses. *Virology* 66, 616-619.
- Lo K, Landau NR, Smale ST. *Mol. Cell. Biol.* 11:5229-5243(1991)
- Maniatis T, Fritsch EF and Sambrook J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY.
- Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT Mullis KB and Erlich HA. (1987). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239, 487-491.
- Sambrook J, Fritsch EF, and Maniatis T. (1989). *Molecular Cloning a Laboratory Manual*, 2nd ed. Cold Spring Harbour Laboratory, Cold Spring Harbour New York.
- Smith GE, Summers MD and Fraser MJ. (1983). Production of human beta interferon in insect cells infected with a baculovirus vector. *Mol. Cell. Biol.*, 3, 2156-2165.
- Stranstrom H, Verjalainen P, Meoning V, Hunsmann G, Schwarz H. and Schafer W. (1974). C- type particles produced by a permanent cell line from a leukemic pig. 1 Origin and properties of host cells and some evidence for the occurrence of C-Type like particles *Virology* 57, 175-178.
- Todaro GJ, Benveniste RE, Lieber MM and Sherr CJ. (1974). Characterizaation of a type C virus released from the porcine cell line PK (15). *Virology* 58, 65-74.

CLAIMS

1. An isolated polynucleotide fragment as shown in Figures 1, 2, 3 or 4, subsequence thereof, or corresponding RNA sequence thereof:

- (a) encoding at least one porcine retrovirus (PoEV) expression product;
- (b) encoding a derivative of said expression product displaying a physiological and/or immunological activity substantially similar to the physiological and/or immunological activity of said expression product as described in (a); or
- (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

2. An isolated polynucleotide fragment according to claim 1:

- (a) encoding at least one polypeptide having an amino acid sequence which is shown in Figures 3 or 4;
- (b) encoding a derivative of said at least one polypeptide displaying a physiological and/or immunological activity substantially similar to the physiological and/or immunological activity of said expression product as described in (a); or
- (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

3. An isolated polynucleotide fragment according to claim 1 or 2:

- (a) encoding the polymerase (POL) polypeptide;
- (b) encoding a derivative polypeptide displaying a physiological and/or immunological activity substantially similar to the physiological and/or immunological activity of the polymerase (POL) polypeptide as described in (a); or
- (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

AMENDED SHEET

037453-02099

4. An isolated polynucleotide fragment according to claim 2:

- (a) encoding the virion core polypeptide (GAG) and/or envelope polypeptide (ENV);
- (b) encoding a derivative polypeptide displaying a physiological and/or immunological activity of said virion core polypeptide (GAG) and/or envelope polypeptide (ENV) as described in (a); or
- (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

5. An isolated polynucleotide fragment displaying at least 90% sequence identity with the sequence as shown in Figures 2 or 3:

- (a) encoding the virion core polypeptide (GAG), polymerase (POL) and envelope polypeptide (ENV) of porcine retrovirus (PoEV);
- (b) encoding a derivative polypeptide displaying a physiological and/or immunological activity substantially similar to the physiological and/or immunological activity of said GAG POL and ENV polypeptides as described in (a); or
- (c) which is complementary to a polynucleotide sequence as defined in (a) or (b).

6. A recombinant nucleic acid molecule comprising a polynucleotide fragment according to any one of claims 1 to 5.

7. A recombinant nucleic acid molecule according to claim 6 wherein the recombinant nucleic acid molecule comprises regulatory control sequences operably linked to said polynucleotide fragment for controlling expression of said polynucleotide fragment.

8. A vector comprising a recombinant nucleic acid molecule according to either of claims 6 or 7.

9. A vector according to claim 8 which is a virus or a plasmid.

10. A prokaryotic or eukaryotic host cell transformed by a polynucleotide fragment, recombinant nucleic acid molecule, or vector according to any of claims 1 to 9.

11. A recombinant PoEV polypeptide comprising an amino acid sequence as shown in Figure 3 displaying POL activity.

12. A recombinant PoEV polypeptide comprising an amino acid sequence with at least 95% sequence identity to the GAG amino acid sequence as shown in Figure 3.

13. A recombinant PoEV polypeptide comprising an amino acid sequence with at least 75% sequence identity to the ENV amino acid sequence as shown in Figures 3 or 4.

14. A recombinant PoEV polypeptide comprising a sequence as shown in Figures 3 or 4, or derivative polypeptide displaying a physiological and/or immunological activity of the PoEV polypeptide.

15. A vaccine comprising a recombinant PoEV polypeptide according to any one of claims 11 to 14, or an inactivated PoEV virus and a pharmaceutically acceptable carrier.

16. An anti-PoEV antibody or fragment thereof raised against a polypeptide or derivative according to any one of claims 11 to 14.

17. A polynucleotide primer which is capable of specifically hybridising to a PoEV polynucleotide fragment as shown in Figures 1, 2, 3 or 4 and capable of initiating chain extension from the 3' end of the primer, but which is not able to correctly initiate chain extension from non PoEV sequences.

AMENDED SHEET

18. A polynucleotide probe which is capable of specifically hybridising under stringent conditions to a polynucleotide sequence as shown in Figures 1, 2, 3 or 4, but not to non PoEV sequences under stringent conditions.

19. A probe or a primer according to claims 17 or 18 which have substantial nucleotide sequence identity with a strand of the molecule depicted in Figures 1, 2, 3 or 4 or a strand complementary therewith, with a corresponding RNA molecule, or with a part of such a molecule.

20. A PoEV detection kit comprising a polynucleotide primer or probe according to any of claims 17 to 19.

21. Use of a PoEV specific polynucleotide primer or probe according to any of claims 17 to 19 in the detection of PoEV in a sample.

22. Use of a PoEV specific polynucleotide primer or primers according to either of claims 17 or 19 in a polymerase chain reaction for the detection of PoEV in a sample.

23. Porcine embryos, embryonic stem cells or cells containing totipotent nuclei capable of forming a viable embryo which have been manipulated by use of a polynucleotide sequence derived from the polynucleotide sequence shown in Figures 1, 2, 3 or 4 so as to not express an infectious, PoEV.

24. A pig obtainable from the porcine embryos, embryonic stem cells or cells containing totipotent nuclei capable of forming a viable embryo according to claim 23.

25. Cells, tissues or organs obtainable from a pig according to claim 24.

26. A recombinant PoEV polypeptide according to any one of claims 11 to 14 for use in the preparation of a vaccine.

27. Use of a polynucleotide primer or probe according to any of claims 17 to 19 in the preparation of a detection kit capable of detecting the presence of PoEV nucleic acid in a sample.

28. A polynucleotide; polypeptide; cells, tissues or organs according to any one of claims 1 to 5, 11 to 13 or 25 for use in therapy or diagnosis.

29. Use of a polynucleotide; polypeptide; cells, tissues or organs according to any one of claims 1 to 5, 11 to 13 or 25 in the preparation of a medicament for therapy or diagnosis.

30. The invention substantially as hereinbefore described.

1/22

Figure 1

1 GAATTCGCGGCCCGCTCGACAGATGCCTTCTTCTGCCTGAGATTACACCCCACTAGCCAA 60
 61 CCACTTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGG 120
 121 ACCCGACTGCCCCAAGGGTCAAGAAGTCCCGACCATCTTTGACGAAGCCCTACACAGG 180
 181 GACCTGGCCAACTTCAGGATCCAACACCCCTCAGGTGACCCCTCCTCCAGTACGTGGATGAC 240
 241 CTGCTTCTGGCGGGAGCCACCAAACAGGACTGCTTAGAAGGTACGAAGGCACTACTGCTG 300
 301 GAATTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTTGCAGGAGAGAG 360
 361 GTAACATACTTGGGGTACAGTTTGCGGGGCGGGCAGCGATGGCTGACGGAGGCACGGAAG 420
 421 AAAACTGTAGTCCAGATACCGGCCCCAACCACAGCCAAACAAGTGAGAGAGTTTTTGGGG 480
 481 ACAGCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGCAGCCTTAGCAGCCCCACTCTAC 540
 541 CCGCTAACCAGAAAAGGGGATTCTCCTGGGCTCCTGAGCACCAGAAGGCATTTGAT 600
 601 GCTATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAATAAACC 660
 661 TTTACCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTAAACCCAAACCCTA 720
 721 GGACCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTGTAGCCAGTGGT 780
 781 TGGCCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGGACGCTGACAAA 840
 841 TTGACTTTGGGACAGAATATAACTGTAATAGCCCCCATGCATTGGAGAACATCGTTCGG 900
 901 CAGCCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAAGCCTGCTTCTC 960
 961 ACAGAGAGGGTCACTTTGCTCCACCAGCCGCTCTCAACCCTGCCACTCTTCTGCCTGAA 1020
 1021 GAGACTGATGAACCAGTGAATCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTC 1080
 1081 CGCAAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCTAACCTGGTTCAGTGACGGA 1140
 1141 AGCAGCTATGTGGTGGAAGGTAAGAGGATGGCTGGGGCGGCAGTGGTGGACGGGACCCGC 1200
 1201 ACGATCTGGGCCAGCAGCCTGCCGGAAGGAACTTCAGCGCAAAGGCTGAGCTCATGGCC 1260

Figure 1 cont.

1261 CTCACGCAAGCTTTGCGGCTGGCCGAAGGGAAATCCATAAACATTTATACGGACAGCAGG 1320
1321 TATGCCTTTGCGACTGCACACGTACACGGGGCCATCTATAAACAAAGGGGGTTGCTTACC 1380
1381 TCAGCAGGGAGGGAAATAAAGAACAAAGAGGAAATTCTAAGCCTATTAGAAGCCTTACAT 1440
1441 TTGCCAAAAAGGCTAGCTATTATACACTGTCTTGACATCAGAAAGCCAAAGATCTCATA 1500
1501 TCTAGAGGGAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCCAGGCTGTTAACCTT 1560
1561 CTGCCTATAATAGAAACGCCCAAAGCCCCAGAAGCCAGACGACAGTACACCTTAGAAGAC 1620
1621 TGGCAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGACCTGCTATACC 1680
1681 TCATATGGGAAGGAAATCCTGCCCCACAAAGAAGGGTTAGAATATGTCCAACAGATACAT 1740
1741 CGTCTAACCACCTAGGAATAAACACCTGCAGCAGTTGGTCAGAACATCCCCTTATCAT 1800
1801 GTTCTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGCCCTGCCAGCTG 1860
1861 GTTAATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAAGCCACCCAGGC 1920
1921 GCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAAACAAATATCTA 1980
1981 TTGGTTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTAAGAAAGAGACT 2040
2041 TCAACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTTCCAAGATTTGGAATACCTAAG 2100
2101 GTAATAGGGTCAGACAATGGTCCAGCTTTTCGTTGCCAGGTAAAGTCAGGGACTGGCCAAG 2160
2161 ATATTGGGGATTGATTGGAAACTGCATTGTGCATACAGACCCCAAAGCTCAGGACAGGTA 2220
2221 GAGAGGATGAATAGAACCATTAAAGAGACCTTACCAAATTGACCACAGAGACTGGCATT 2280
2281 AATGATTGGATGGCTCTCCTGCCCTTTGTGCTTTTTAGGGTGAGGAACACCCCTGGACAG 2340
2341 TTTGGGCTGACCCCTATGAATTGCTCTACGGGGGACCCCCCGTTGGCAGAAATTGCC 2400
2401 TTTGCACATAGTGCTGATGTGCTGCTTTCCAGCCTTTGTTCTCTAGGCTCAAGGCGCTC 2460
2461 GAGTGGGTGAGGCAGCGAGCGTGGAAGCAGCTCCGGGAGGCCTACTCAGGAGGAGACTTG 2520

Figure 1 cont.

2521 CAAGTTCACATCGCTTCCAAGTTGGAGATTCAGTCTATGTTAGACGCCACCGTGCAGGA 2580
2581 AACCTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTTGACCACACCAACGGCTGTG 2640
2641 AAAGTCGAAGGAATCCCCACCTGGATCCATGCATCCCACGTTAAGCCGGCGCCACCTCCC 2700
2701 GATTCGGGGTGGAAAGCCGAAAAGACTGAAAATCCCCTTAAGCTTCGCCTCCATCGCGTG 2760
2761 GTTCCTTACTCTGTCAATAACTCCTCAAGTTAATGGTAAACGCCTTGTGGACAGCCCGAA 2820
2821 CTCCCATAAACCCTTATCTCTCACCTGGTTACTTACTGACTCCGGTACAGGTATTAATAT 2880
2881 TAACAGCACTCAAGGGGAGGCTCCCTTGGGGACCTGGTGGCCTGAATTATATGTCTGCCT 2940
2941 TCGATCAGTAATCCCTGGTCTCAATGACCAGGCCACACCCCCGATGTACTCCGTGCTTA 3000
3001 CGGGTTTTACGTTTGCCAGGACCCCCAAATAATGAAGAATATTGTGGAAATCCTCAGGA 3060
3061 TTTCCTTTGCAAGCAATGGAGCTGCATAACTTCTAATGATGGGAATTGGAAATGGCCAGT 3120
3121 CTCTCAGCAAGACAGAGTAAGTTACTCTTTTGTTAACAATCCTACCAGTTATAATCAATT 3180
3181 TAATTATGGCCATGGGAGATGGAAAGATTGGCAACAGCGGGTACAAAAAGATGTACGAAA 3240
3241 TAAGCAAATAAGCTGTCATTGTTAGACCTAGATTACTTAAAAATAAGTTTCACTAAAAA 3300
3301 AAAAAAAAAAAAAAAAAAAAAA 3320

4/22

Figure 2

1 TGTGGGCCCCAGCGCGCTTGGAATAAAAATCCTCTTGCTGTTTGCATCAAGACCGCTTCT 60
61 CGTGAGTGATTGGGGTGTGCGCTCTTCCGAGCCCGGACGAGGGGGATTGTTCTTTTACT 120
121 GGCCTTTCATTTGGTGCCTTGGCCGGGAAATCCTGCGACCACCCCTTACACCCGAGAACC 180
181 GACTTGGAGGTAAAGGGATCCCTTTTGAACGTGTGTGTGTGTGCGGCCGGCGTCTCTGTT 240
241 CTGAGTGTCTGTTTTCGGTGATGCGCGCTTTCGGTTTGCAGCTGTCTCTCAGACCGTAA 300
301 GGA CTGAGGACTGTGATCAGCAGACGTGCTAGGAGGATCACAGGCTGCCACCCTGGGGG 360
361 ACGCCCCGGGAGGTGGGAGAGCCAGGACGCGCTGGTGGTCTCCTACTGTGGT CAGAGG 420
421 ACCGAGTTCTGTTGTTGAAGCGAAAGCTTCCCCCTCCGCGGCCGTCCGACTCTTTTGCCT 480
481 GCTTGTGGAAGACGCGGACGGGTGCGGTGTGTCTGGATCTGTTGGTTTCTGTCTCGTGTG 540
541 TCTTTGTCTTGTGCGTCCTTGTCTACAGTTTTAATATGGGACAGACAGTGACTACCCCC 600
601 TTAGTTTGACTCTCGACCATTGGACTGAAGTTAGATCCAGGGCTCATAATTTGTCAGTTC 660
661 AGGTTAAGAAGGGACCTTGGCAGACTTCTGTGCCTCTGAATGGCCAACATTGATGTTG 720
721 GATGGCCATCAGAGGGGACCTTTAATTCTGAAATTATCCTGGCTGTTAAGGCAATCATT 780
781 TTCAGACTGGACCCGGCTCTCATCCTGATCAGGAGCCCTATATCCTTACGTGGCAAGATT 840
841 TGGCAGAAGATCCTCCGCCATGGGTAAACCATGGCTAAATAAACCAAGAAAGCCAGGTC 900
901 CCCGAATCCTGGCTCTTGGAGAGAAAAACAAACTCGGCCGAAAAAGTCGAGCCCTCTT 960
961 CCTCGTATCTACCCCGAGATCGAGGAGCCGCCGACTTGGCCGGAACCCCAACCTGTTCCC 1020
1021 CCACCCCTTATCCAGCACAGGTGCTGTGAGGGGACCTCTGCCCCCTCTGGAGCTCCGG 1080
1081 TGGTGGAGGGACCTGCTGCCGGGACTCGGAGCCGGAGAGGCGCCACCCCGGAGCGGACAG 1140
1141 ACGAGATCGGATATTACCGCTGCGCACCTATGGCCCTCCCATGCCAGGGGGCCAATTGC 1200
1201 AGCCCTCCAGTATTGGCCCTTTTCTTCTGCAGATCTCTATAATTGGAAAATAACCATC 1260

5/22

Figure 2 cont.

1261 CCCCTTTCTCGGAGGATCCCAACGCCTCACGGGGTTGGTGGAGTCCCTTATGTTCTCTC 1320
1321 ACCAGCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTCACAACCGAGGAGC 1380
1381 GAGAGAGAATTCTGTTAGAGGCTAGAAAAATGTTCTGGGGCCGACGGGCGACCCACGC 1440
1441 AGTTGCAAAATGAGATTGACATGGGATTTCCCTTGACTCGCCCCGGTTGGGACTACAACA 1500
1501 CGGCTGAAGGTAGGGAGAGCTTGAAAACTATCGCCAGGCTCTGGTGGCGGGTCTCCGGG 1560
1561 GCGCCTCAAGACGGCCCACTAATTTGGCTAAGGTAAGAGAGGTGATGCAGGGACCGAACG 1620
1621 AACCTCCCTCGGTATTTCTTGAGAGGCTCATGGAAGCCTTCAGGCGGTTCACCCCTTTTG 1680
1681 ATCTACCTCAGAGGCCCAAGAAAGCCTCAGTGGCCCTGGCCTTCATTGGGCAGTCGGCTC 1740
1741 TGGATATCAGGAAGAACTTCAGAGACTGGAAGGGTTACAGGAGGCTGAGTTACGTGATC 1800
1801 TAGTGAGAGAGGCAGAGAAGGTGTATTACAGAAGGGAGACAGAAGAGGAGAAGGAACAGA 1860
1861 GAAAAGAAAAGGAGAGAGAAGAAAGGGAGGAAAGACGTGATAGACGGCAAGAGAAGAATT 1920
1921 TGAATAAGATCTTGGCCGCAGTGGTTGAAGGGAAGAGCAGCAGGGAGAGAGAGAGAGATT 1980
1981 TTAGGAAAATTAGGTCAGGCCCTAGACAGTCAGGGAACCTGGGCAATAGGACCCCACTCG 2040
2041 ACAAGGACCAGTGTGCGTATTGTAAAGAAAAAGGACACTGGGCAAGGAACTGCCCCAAGA 2100
2101 AGGGAAACAAAGGACCGAAGTCCTAGCTCTAGAAGAAGATAAAGATTAGGGGAGACGGGT 2160
2161 TCGGACCCCCTCCCCGAGCCAGGGTAACCTTTGAAGGTGGAGGGGCAACCAGTTGAGTTC 2220
2221 CTGGTTGATACCGGAGCGGAGCATTCACTGCTGCTACAACCATTAGGAAAATAAAAGAA 2280
2281 AAAAAATCCTGGGTGATGGGTGCCACAGGGCAACGGCAGTATCCATGGACTACCCGAAGA 2340
2341 ACCGTTGACTTGGGAGTGGGACGGGTAACCCACTCGTTTCTGGTCATCCCTGAGTGCCCA 2400
2401 GTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAATTTCTTTTGAACAA 2460
2461 GGAAGACCAGAAGTGTCTGTGAATAACAAACCCATCACTGTGTTGACCCTCCAATTAGAT 2520

6/22

Figure 2 cont.

2521 GATGAATATCGACTATATTCTCCCCAAGTAAAGCCTGATCAAGATATACAGTCCTGGTTG 2580

2581 GAGCAGTTTCCCCAAGCCTGGGCAGAAACCGCAGGGATGGGTTTGGCAAAGCAAGTTCCC 2640

2641 CCACAGGTTATTCAACTGAAGGCCAGTGCTACACCAGTATCAGTCAGACAGTACCCCTTG 2700

2701 AGTAGAGAGGCTCGAGAAGGAATTTGGCCGCATGTTCAAAGATTAATCCAACAGGGCATC 2760

2761 CTAGTTCCTGTCCAATCCCCTTGGAACTACTCCCCTGCTACCGGTTAGGAAGCCTGGGACC 2820

2821 AATGATTATCGACCAGTACAGGACTTGAGAGAGGTCAATAAAAGGGTGCAGGACATACAC 2880

2881 CCAACGGTCCCGAACCCTTATAACCTCTTGAGCGCCCTCCCGCCTGAACGGAAGTGGTAC 2940

2941 ACAGTATTGGACTTAAAAGATGCCTTCTTCTGCCTGAGATTACACCCCACTAGCCAACCA 3000

3001 CTTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCGGGCAGCTCACCTGGACC 3060

3061 CGACTGCCCCAAGGGTCAAGAAGTCCCGGACCATCTTGACGAAGCCCTACACAGGGAC 3120

3121 CTGGCCAACTTCAGGATCCAACACCCCTCAGGTGACCCTCCTCCAGTACGTGGATGACCTG 3180

3181 CTTCTGGCGGGAGCCACCAACAGGACTGCTTAGAAGGTACGAAGGCACTACTGCTGGAA 3240

3241 TTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTGTCAGGAGAGAGGTA 3300

3301 ACATACTTGGGGTACAGTTTGCGGGGCGGGCAGCGATGGCTGACGGAGGCACGGAAGAAA 3360

3361 ACTGTAGTCCAGATACCGGCCCCCAACCACAGCCAAACAAGTGAGAGAGTTTTTGGGGACA 3420

3421 GCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGCAGCCTTAGCAGCCCCACTCTACCCG 3480

3481 CTAACCAAAGAAAAGGGGGATTCTCCTGGGCTCCTGAGCACCAGAAGGCATTTGATGCT 3540

3541 ATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACGTAACCTAAACCCTTT 3600

3601 ACCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCGAGGAGTTTAAACCCAAACCCTAGGA 3660

3661 CCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTGTAGCCAGTGGTTGG 3720

3721 CCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGGACGCTGACAAATTG 3780

7/22

Figure 2 cont.

3781 ACTTTGGGACAGAATATAACTGTAATAGCCCCCATGCATTGGAGAACATCGTTCGGCAG 3840

3841 CCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAAGCCTGCTTCTCACA 3900

3901 GAGAGGGTCACTTTGCTCCACCAGCCGCTCTCAACCCCTGCCACTCTTCTGCCTGAAGAG 3960

3961 ACTGATGAACCACTGACTCATGATTGCCATCAACTATTGATTGAGGAGACTGGGGTCCGC 4020

4021 AAGGACCTTACAGACATAACCGCTGACTGGAGAAGTGCTAACCTGGTTCACTGACGGAAGC 4080

4081 AGCTATGTGGTGAAGGTAAGAGGATGGCTGGGGCGGCAGTGGTGGACGGGACCCGCACG 4140

4141 ATCTGGGCCAGCAGCCTGCCGGAAGGAACTTCAGCGCAAAAGGCTGAGCTCATGGCCCTC 4200

4201 ACGCAAGCTTTGCGGCTGGCCGAAGGGAAATCCATAAACATTTATACGGACAGCAGGTAT 4260

4261 GCCTTTGCGACTGCACACGTACACGGGGCCATCTATAAACAAAGGGGGTTGCTTACCTCA 4320

4321 GCAGGGAGGGAAATAAAGAACAAAGAGGAAATCTAAGCCTATTAGAAGCCTTACATTTG 4380

4381 CCAAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCAAAGATCTCATATCT 4440

4441 AGAGGGAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCCAGGCTGTAAACCTTCTG 4500

4501 CCTATAATAGAAACGCCCAAAGCCCCAGAACCCAGACGACAGTACACCCTAGAAGACTGG 4560

4561 CAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGACCTGCTATACCTCA 4620

4621 TATGGGAAGGAAATCCTGCCCCACAAAGAAGGGTTAGAATATGTCCAACAGATACATCGT 4680

4681 CTAACCCACCTAGGAACCTAAACACCTGCAGCAGTTGGTCAGAACATCCCCTTATCATGTT 4740

4741 CTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGCCCTGCCAGCTGGTT 4800

4801 AATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAAGCCACCCAGGCGCT 4860

4861 CACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAAACAAATATCTATTG 4920

4921 GTTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTAAGAAAGAGACTTCA 4980

4981 ACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTTCCAAGATTGGAATACCTAAGGTA 5040

8/22

Figure 2 cont.

5041 ATAGGGTCAGACAATGGTCCAGCTTTCGTTGCCAGGTAAGTCAGGGACTGGCCAAGATA 5100

5101 TTGGGGATTGATTGGAACTGCATTGTGCATACAGACCCCAAAGCTCAGGACAGGTAGAG 5160

5161 AGGATGAATAGAACCATTAAAGAGACCCCTTACCAAATTGACCACAGAGACTGGCATTAAAT 5220

5221 GATTGGATGGCTCTCCTGCCCTTTGTGCTTTTATAGGGTGAGGAACACCCCTGGACAGTTT 5280

5281 GGGCTGACCCCTATGAATTGCTCTACGGGGGACCCCCCGTTGGCAGAAATTGCCTTT 5340

5341 GCACATAGTGCTGATGTGCTGCTTTCCAGCCTTTGTTCTCTAGGCTCAAGGCGCTCGAG 5400

5401 TGGGTGAGGCAGCGAGCGTGGAAGCAGCTCCGGGAGGCCTACTCAGGAGGAGACTTGCAA 5460

5461 GTTCCACATCGCTTCCAAGTTGGAGATTCAGTCTATGTTAGACGCCACCGTGCAGGAAAC 5520

5521 CTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTTGACCACACCAACGGCTGTGAAA 5580

5581 GTCGAAGGAATCCCCACCTGGATCCATGCATCCACGTTAAGCYGGCGCCACCTCCCGAC 5640

5641 TCGGGGTGGAGAGCCGAAAAGACTGAGAATCCCTTAAGCTTCGCCTCCATCGCCTGGTT 5700

5701 CCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCCTTATAGACAGCTCGAACCC 5760

5761 CCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCTGATACGGGTGTCACTGTAAA 5820

5821 TAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAAGTGCATTTCTGCCTCCG 5880

5881 ATTGATTAAACCCGCTGTTAARAGCACACCTCCCAACCTAGTCCGTAGTTATGGGTTCTA 5940

5941 TTGCTGCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTCTGGGGAATCCTTCTGTAG 6000

6001 GAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCGATCTCTCTCCAGGA 6060

6061 CCGGGTAAATTTCTCTTTGTCAATTCCGGCCCGGGCAAGTACAAAATGATGAACTATA 6120

6121 TAAAGATAAGAGCTGCTCCCATCAGACTTAGATTATCTAAAGATAAGTTTCACTGAAAG 6180

6181 GAAAACAGGAAAATATTCAAAAGTGGATAAATGGTATGAGCTGGGGAATAGTTTTTTATT 6240

6241 ATATGGCGGGGAGCAGGTCCTTAAACCATTGCGCTTAGGATAGAGACGGGGACAGA 6300

9/22

Figure 2 cont.

6301 ACCCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGGCCCCCGGCCCTGGA 6360
6361 GCCACCGCATAACTTGCCGGTGCCCCAATTAACCTCGCTGCGGCCTGACATAACACAGCC 6420
6421 GCCTAGCAACAGTACCACTGGATTGATTCTACCAACACGCCTAGAACTCCCCAGGTGT 6480
6481 TCCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCTTCCAAGCCATCAA 6540
6541 CTCCACCGACCCTGATGCCACTTCTTCTTGTGGCTTTGTCTATCCTCAGGGCCTCCTTA 6600
6601 TTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAGCATAGAAATCAATG 6660
6661 TACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGGAAGGGGACATGCAT 6720
6721 AGGAAAAGCTCCCCCATCCACCAACACCTTTGCTATAGTACTGTGGTTTATGAGCAGGC 6780
6781 CTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGGCATGCAATACTGGGTT 6840
6841 AACCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTCTGTGTCATGGTCCA 6900
6901 AATCGTCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGATGAATATGACTATCG 6960
6961 GTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTAATGCTCGGATTAGG 7020
7021 GACGGCCGTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGACCACAGCAGCTAGA 7080
7081 GAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGAGCCTTAAAGGAGTC 7140
7141 TGTTAGCAACCTAGAAGAGTCCCTGACTTCTTTGTCTGAAGTGGTTCTACAGAACCGGAG 7200
7201 GGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGGTTATGTGCAGCCTTAAAAGAAGAATG 7260
7261 TTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAACAAGCTTAGAAAAAA 7320
7321 GTTAGAGAGGCGTCGAAGGGAAAGAGAGGCTGACCAGGGGTGGTTTGAAGGATGGTTCAA 7380
7381 CAGGTCTCCTTGGATGACCACCCTGCTTTCTGCTCTGACGGGGCCCCCTAGTAGTCCTGCT 7440
7441 CCTGTTACTTACAGTTGGGCCTTGCTTAATTAATAGGTTTGTTCCTTTGTTAGAGAACG 7500
7501 AGTGAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGCCTTCTGAGCCAAGG 7560

10/22

Figure 2 cont.

7561 AGAAACTGACCTCTAGCCTTCCCAGTTCTAAGATTAGAACTATTAACAAGACAAGAAGTG 7620
7621 GGAATGAAAGGATGAAAATGCAACCTAACCTCCCAGAAGCCAGGAAGTTAATAAAAAG 7680
7681 CTCTAAATGCCCCGAATTMCAGACCCTGCTGGCTGCCAGTAAATAGGTAGAAGGTCACA 7740
7741 CTTCTATTGTTCCAGGGCCTGCTATCCTGGCCTAAGTAAGATAACAGGAAATGAGTTGA 7800
7801 CTAATCGCTTATCTGGATTCTGTAAACTGACTGGCACCATAGAAGAATTGATTACACAT 7860
7861 TGACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCAGGAGCCCACGCAG 7920
7921 ATGCGGACCTCCGGAGCTATTTTAAAATGATTGGTCCACGGAGCGCGGGCTCTCGATATT 7980
7981 TTAATAATGATTGGTCCATGGAGCGCGGGCTCTCGATATTTTAAAATGATTGGTTTGTGAC 8040
8041 GCACAGGCTTTGTTGTGAACCCCATAAAAGCTGTCCCGATTCCGCACTCGGGGCCGCAGT 8100
8101 CCTCTACCCCTGCGTGGTGTACGACTGTGGGCCCCAGCGCGCTTGAATAAAAATCCTCT 8160
8161 TGCTGTTTGCATCAAAAAAAAAAAAAAAAAAAAAA 8196

11/22

Figure 3

1 GTGGTGTACGACTGTGGGCCCCAGCGCGCTTGAATAAAAAATCCTCTTGCTGTTTGCATC 60
 61 AAGACCGCTTCTCGTGAGTGATTTGGGGTGTGCGCTCTTCCGAGCCCGGACGAGGGGGAT 120
 121 TGTTCCTTTTACTGGCCTTTTCATTTGGTGCCTTGGCCGGGAAATCCTGCGACCACTCCCTTA 180
 181 CACCCGAGAACCGACTTGGAGGTAAAGGGATCCCCTTTGGAACGTGTGTGTGTGTGCGGCC 240
 241 GCGTCTCTGTTCTGAGTGCTGTTTTCGGTGATGCGCGCTTTCGGTTTGCAGCTGTCTT 300
 301 CTCAGACCGTAAGGACTGGAGGACTGTGATCAGCAGACGTGCTAGGAGGATCACAGGCTG 360
 361 CCACCCCTGGGGGACGCCCCGGGAGGTGGGGAGAGCCAGGGACGCCTGGTGGTCTCCTACT 420
 421 GTCGGTCAGAGGACCGAGTTCTGTTGTTGAAGCGAAAGCTTCCCCCTCCGCGGCCGTCCG 480
 481 ACTCTTTTGCCTGCTTGTGGAAGACGCGGACGGGTGCGGTGTGTCTGGATCTGTTGGTTT 540
 541 CTGTCTCGTGTGCTTTGTCTTGTGCGTCTTGTCTACAGTTTTAATATGGGACAGACAG 600
 MetGlyGlnThrV
 601 TGACTACCCCTTAGTTTGAATCTCGACCATTTGGACTGAAGTTAGATCCAGGGCTCATA 660
 alThrThrProLeuSerLeuThrLeuAspHisTrpThrGluValArgSerArgAlaHisA
 661 ATTTGTCAGTTCAGGTTAAGAAGGGACCTTGGCAGACTTTCTGTGCCTCTGAATGGCCAA 720
 snLeuSerValGlnValLysLysGlyProTrpGlnThrPheCysAlaSerGluTrpProT
 721 CATTCGATGTTGGATGGCCATCAGAGGGGACCTTTAATTCTGAAATTATCCTGGCTGTTA 780
 hrPheAspValGlyTrpProSerGluGlyThrPheAsnSerGluIleIleLeuAlaValL
 781 AGGCAATCATTTTTTCAGACTGGACCCGGCTCTCATCCTGATCAGGAGCCCTATATCCTTA 840
 ysAlaIleIlePheGlnThrGlyProGlySerHisProAspGlnGluProTyrIleLeuT
 841 CGTGGCAAGATTGGCAGAAGATCCTCCGCCATGGGTAAACCATGGCTAAATAAACCAA 900
 hrTrpGlnAspLeuAlaGluAspProProProTrpValLysProTrpLeuAsnLysProA
 901 GAAAGCCAGGTCCCGAATCCTGGCTCTTGGAGAGAAAAACAAACACTCGGCCGAAAAAG 960
 rgLysProGlyProArgIleLeuAlaLeuGlyGluLysAsnLysHisSerAlaGluLysV
 961 TCGAGCCCTCTTCCTCGTATCTACCCCGAGATCGAGGAGCCGCGACTTGGCCGGAACCC 1020
 alGluProSerSerSerTyrLeuProArgAspArgGlyAlaAlaAspLeuAlaGlyThrP
 1021 CAACCTGTTCCCCACCCCTTATCCAGCACAGGGTGTGTGAGGGGACCTCTGCCCTC 1080
 roThrCysSerProThrProLeuSerSerThrGlyCysCysGluGlyThrSerAlaProP

12/22

Figure 3 cont.

| | | |
|------|---|------|
| 1081 | CTGGAGCTCCGGTGGTGGAGGGACCTGCTGCCGGGACTCGGAGCCGGAGAGGCCACCC | 1140 |
| | roGlyAlaProValValGluGlyProAlaAlaGlyThrArgSerArgArgGlyAlaThrP | |
| 1141 | CGGAGCGGACAGACGAGATCGCGATATTACCGCTGCGCACCTATGGCCCTCCCATGCCAG | 1200 |
| | roGluArgThrAspGluIleAlaIleLeuProLeuArgThrTyrGlyProProMetProG | |
| 1201 | GGGGCCAATTGCAGCCCCTCCAGTATTGGCCCTTTTCTTCTGCAGATCTCTATAATTGGA | 1260 |
| | lyGlyGlnLeuGlnProLeuGlnTyrTrpProPheSerSerAlaAspLeuTyrAsnTrpL | |
| 1261 | AAACTAACCATCCCCCTTTCTCGGAGGATCCCCAACGCCTCACGGGGTTGGTGGAGTCCC | 1320 |
| | ysThrAsnHisProProPheSerGluAspProGlnArgLeuThrGlyLeuValGluSerL | |
| 1321 | TTATGTTCTCTCACCAGCCTACTTGGGATGATTGTCAACAGCTGCTGCAGACACTCTTCA | 1380 |
| | euMetPheSerHisGlnProThrTrpAspAspCysGlnGlnLeuLeuGlnThrLeuPheT | |
| 1381 | CAACCGAGGAGCGAGAGAGAATTCTGTTAGAGGCTAGAAAAAATGTTCTGGGGCCGACG | 1440 |
| | hrThrGluGluArgGluArgIleLeuLeuGluAlaArgLysAsnValProGlyAlaAspG | |
| 1441 | GGCGACCCACGCAGTTGCAAAATGAGATTGACATGGGATTTCCCTTGACTCGCCCCGGTT | 1500 |
| | lyArgProThrGlnLeuGlnAsnGluIleAspMetGlyPheProLeuThrArgProGlyT | |
| 1501 | GGGACTACAACACGGCTGAAGGTAGGGAGAGCTTGAAAATCTATCGCCAGGCTCTGGTGG | 1560 |
| | rpAspTyrAsnThrAlaGluGlyArgGluSerLeuLysIleTyrArgGlnAlaLeuValA | |
| 1561 | CGGGTCTCCGGGGCGCCTCAAGACGGCCCCACTAATTTGGCTAAGGTAAGAGAGGTGATGC | 1620 |
| | laGlyLeuArgGlyAlaSerArgArgProThrAsnLeuAlaLysValArgGluValMetG | |
| 1621 | AGGGACCGAACGAACCTCCCTCGGTATTTCTTGAGAGGCTCATGGAAGCCTTCAGGCGGT | 1680 |
| | lnGlyProAsnGluProProSerValPheLeuGluArgLeuMetGluAlaPheArgArgP | |
| 1681 | TCACCCCTTTTGATCCTACCTCAGAGGCCAGAAAGCCTCAGTGGCCCTGGCCTTCATTG | 1740 |
| | heThrProPheAspProThrSerGluAlaGlnLysAlaSerValAlaLeuAlaPheIleG | |
| 1741 | GGCAGTCGGCTCTGGATATCAGGAAGAACTTCAGAGACTGGAAGGGTTACAGGAGGCTG | 1800 |
| | lyGlnSerAlaLeuAspIleArgLysLysLeuGlnArgLeuGluGlyLeuGlnGluAlaG | |
| 1801 | AGTTACGTGATCTAGTGAGAGAGGCAGAGAAGGTGTATTACAGAAGGGAGACAGAAGAGG | 1860 |
| | luLeuArgAspLeuValArgGluAlaGluLysValTyrTyrArgArgGluThrGluGluG | |
| 1861 | AGAAGGAACAGAGAAAAGAAAAGGAGAGAGAAGAAAGGGAGGAAAGACGTGATAGACGGC | 1920 |
| | luLysGluGlnArgLysGluLysGluArgGluGluArgGluGluArgArgAspArgArgG | |
| 1921 | AAGAGAAGAATTTGACTAAGATCTTGGCCGAGTGGTTGAAGGGAAGAGCAGCAGGGAGA | 1980 |
| | lnGluLysAsnLeuThrLysIleLeuAlaAlaValValGluGlyLysSerSerArgGluA | |
| 1981 | GAGAGAGAGATTTTAGGAAAATTAGGTCAGGCCCTAGACAGTCAGGGAACCTGGGCAATA | 2040 |
| | rgGluArgAspPheArgLysIleArgSerGlyProArgGlnSerGlyAsnLeuGlyAsnA | |

Figure 3 cont

2041 GGACCCCACTCGACAAGGACCAGTGTGCGTATTGTAAAGAAAAAGGACACTGGGCAAGGA 2100
rgThrProLeuAspLysAspGlnCysAlaTyrCysLysGluLysGlyHisTrpAlaArgA

2101 ACTGCCCCAAGAAGGGAAACAAAGGACCGAAGgTCCTAGCTCTAGAAGAAGATAAAGATT 2160
snCysProLysLysGlyAsnLysGlyProLysValLeuAlaLeuGluGluAspLysAspE

2161 AGGGGAGACGGGgTTCGGACCCCTCCCCGAGCCCAGGGTAACTTTGAAGGTGGAGGGGC 2220
ndGlyArgArgGlySerAspProLeuProGluProArgValThrLeuLysValGluGlyG

2221 AACCAGTTGAGTTCCTGGTTGATACCGGAGCGGAGCATTGCTGCTGCTACAACCATTAG 2280
lnProValGluPheLeuValAspThrGlyAlaGluHisSerValLeuLeuGlnProLeuG

2281 GAAACTAAAAGAAAAAAATCCTGGGTGATGGGTGCCACAGGGCAACGGCAGTATCCAT 2340
lyLysLeuLysGluLysLysSerTrpValMetGlyAlaThrGlyGlnArgGlnTyrProT

2341 GGACTACCCGAAGAACCGTTGACTTGGGAGTGGGACGGGTAAACCCACTCGTTTCTGGTCA 2400
rpThrThrArgArgThrValAspLeuGlyValGlyArgValThrHisSerPheLeuValI

2401 TCCCTGAGTGCCcAGTACCCCTTCTAGGTAGAGACTTACTGACCAAGATGGGAGCTCAAA 2460
leProGluCysProValProLeuLeuGlyArgAspLeuLeuThrLysMetGlyAlaGlnI

2461 TTTCTTTTGAACAAGGAAGACCAGAAGTGTCTGTGAATAACAAACCCATCACTGTGTTGA 2520
leSerPheGluGlnGlyArgProGluValSerValAsnAsnLysProIleThrValLeuT

2521 CCCTCCAATTAGATGATGAATATCGACTATATTCTCCCCAAGTAAAGCCTGATCAAGATA 2580
hrLeuGlnLeuAspAspGluTyrArgLeuTyrSerProGlnValLysProAspGlnAspI

2581 TACAGTCCTGGTTGGAGCAGTTTCCCCAAGCCTGGGCAGAAACCGCAGGGATGGGTTTGG 2640
leGlnSerTrpLeuGluGlnPheProGlnAlaTrpAlaGluThrAlaGlyMetGlyLeuA

2641 CAAAGCAAGTTCCCCACAGGTTATTCAACTGAAGGCCAGTGCTACACCAGTATCAGTCA 2700
laLysGlnValProProGlnValIleGlnLeuLysAlaSerAlaThrProValSerValA

2701 GACAGTACCCCTTGAGTAGAGAGGCTCGAGAAGGAATTTGGCCGCATGTTCAAAGATTAA 2760
rgGlnTyrProLeuSerArgGluAlaArgGluGlyIleTrpProHisValGlnArgLeuI

2761 TCCAACAGGGCATCCTAGTTCCTGTCCAATCCCCTTGGAACTACTCCCCTGCTACCGGTTA 2820
leGlnGlnGlyIleLeuValProValGlnSerProTrpAsnThrProLeuLeuProValA

2821 GGAAGCCTGGGACCAATGATTATCGACCACTACAGGACTTGAGAGAGGTCAATAAAAGGG 2880
rgLysProGlyThrAsnAspTyrArgProValGlnAspLeuArgGluValAsnLysArgV

2881 TGCAGGACATACACCCAACGGTCCCGAACCCCTTATAACCTCTTGAGCGCCCTCCCGCCTG 2940
alGlnAspIleHisProThrValProAsnProTyrAsnLeuLeuSerAlaLeuProProG

2941 AACGGAACTGGTACACAGTATTGGACTTAAAAGATGCCTTCTTCTGCCTGAGATTACACC 3000
luArgAsnTrpTyrThrValLeuAspLeuLysAspAlaPhePheCysLeuArgLeuHisP

14/22

Figure 3 cont.

| | | |
|------|---|------|
| 3001 | CCACTAGCCAACCACTTTTTGCCTTCGAATGGAGAGATCCAGGTACGGGAAGAACCAGGGC roThrSerGlnProLeuPheAlaPheGluTrpArgAspProGlyThrGlyArgThrGlyG | 3060 |
| 3061 | AGCTCACCTGGACCCGACTGCCCCAAGGGTTCAAGAACTCCCCGACCATCTTTGACGAAG lnLeuThrTrpThrArgLeuProGlnGlyPheLysAsnSerProThrIlePheAspGluA | 3120 |
| 3121 | CCCTACACAGGGACCTGGCCAACTTCAGGATCCAACACCCTCAGGTGACCCTCCTCCAGT laLeuHisArgAspLeuAlaAsnPheArgIleGlnHisProGlnValThrLeuLeuGlnT | 3180 |
| 3181 | ACGTGGATGACCTGCTTCTGGCGGGAGCCACCAAACAGGACTGCTTAGAAGGTACGAAGG yrValAspAspLeuLeuLeuAlaGlyAlaThrLysGlnAspCysLeuGluGlyThrLysA | 3240 |
| 3241 | CACTACTGCTGGAATTGTCTGACCTAGGCTACAGAGCCTCTGCTAAGAAGGCCAGATTT laLeuLeuLeuGluLeuSerAspLeuGlyTyrArgAlaSerAlaLysLysAlaGlnIleC | 3300 |
| 3301 | GCAGGAGAGAGGTAACATACTTGGGGTACAGTTTGCGGGGCGGGCAGCGATGGCTGACGG ysArgArgGluValThrTyrLeuGlyTyrSerLeuArgGlyGlyGlnArgTrpLeuThrG | 3360 |
| 3361 | AGGCACGGAAGAAACTGTAGTCCAGATACCGGCCCCAACCACAGCCAAACAAGTGAGAG luAlaArgLysLysThrValValGlnIleProAlaProThrThrAlaLysGlnValArgG | 3420 |
| 3421 | AGTTTTTGGGGACAGCTGGATTTTGCAGACTGTGGATCCCGGGGTTTGGCGACCTTAGCAG luPheLeuGlyThrAlaGlyPheCysArgLeuTrpIleProGlyPheAlaThrLeuAlaA | 3480 |
| 3481 | CCCCACTCTACCCGCTAACCAAAGAAAAAGGGGGATTCTCCTGGGCTCCTGAGCACCAGA laProLeuTyrProLeuThrLysGluLysGlyGlyPheSerTrpAlaProGluHisGlnL | 3540 |
| 3541 | AGGCATTTGATGCTATCAAAAAGGCCCTGCTGAGCGCACCTGCTCTGGCCCTCCCTGACG ysAlaPheAspAlaIleLysLysAlaLeuLeuSerAlaProAlaLeuAlaLeuProAspV | 3600 |
| 3601 | TAACTAAACCCTTTACCCTTTATGTGGATGAGCGTAAGGGAGTAGCCCCAGGAGTTTAA alThrLysProPheThrLeuTyrValAspGluArgLysGlyValAlaArgGlyValLeuT | 3660 |
| 3661 | CCCAAACCCTAGGACCATGGAGGAGACCTGTTGCCTACCTGTCAAAGAAGCTTGATCCTG hrGlnThrLeuGlyProTrpArgArgProValAlaTyrLeuSerLysLysLeuAspProV | 3720 |
| 3721 | TAGCCAGTGGTTGGCCCGTATGTCTGAAGGCTATCGCAGCTGTGGCCATACTGGTCAAGG alAlaSerGlyTrpProValCysLeuLysAlaIleAlaAlaValAlaIleLeuValLysA | 3780 |
| 3781 | ACGCTGACAAATTGACTTTGGGACAGAATATAACTGTAATAGCCCCCATGCATTGGAGA spAlaAspLysLeuThrLeuGlyGlnAsnIleThrValIleAlaProHisAlaLeuGluA | 3840 |
| 3841 | ACATCGTTCGGCAGCCCCCAGACCGATGGATGACCAACGCCCGCATGACCCACTATCAAA snIleValArgGlnProProAspArgTrpMetThrAsnAlaArgMetThrHisTyrGlnS | 3900 |
| 3901 | GCCTGCTTCTCACAGAGAGGGTCACTTTGCTCCACCAGCCGCTCTCAACCCTGCCACTC erLeuLeuLeuThrGluArgValThrPheAlaProProAlaAlaLeuAsnProAlaThrL | 3960 |

Figure 3 cont

3961 TTCTGCCTGAAGAGACTGATGAACCAGTGACTCATGATTGCCATCAACTATTGATTGAGG 4020
euLeuProGluGluThrAspGluProValThrHisAspCysHisGlnLeuLeuIleGluG

4021 AGACTGGGGTCCGCAAGGACCTTACAGACATACCGCTGACTGGAGAAGTGCTAACCTGGT 4080
luThrGlyValArgLysAspLeuThrAspIleProLeuThrGlyGluValLeuThrTrpP

4081 TCACTGACGGAAGCAGCTATGTGGTGGAGSTAAGAGGATGGCTGGGGCGGCAGTGGTGG 4140
heThrAspGlySerSerTyrValValGluGlyLysArgMetAlaGlyAlaAlaValValA

4141 ACGGGACCCGACGATCTGGGCCAGCAGCCTGCCGGAAGGAACTTCAGCGCAAAAGGCTG 4200
spGlyThrArgThrIleTrpAlaSerSerLeuProGluGlyThrSerAlaGlnLysAlaG

4201 AGCTCATGGCCCTCACGCAAGCTTTGCGGCTGGCCGAAGGGAAATCCATAAACATTTATA 4260
luLeuMetAlaLeuThrGlnAlaLeuArgLeuAlaGluGlyLysSerIleAsnIleTyrT

4261 CGGACAGCAGGTATGCCTTTGCGACTGCACACGTACACGGGGCCATCTATAAACAAAGGG 4320
hrAspSerArgTyrAlaPheAlaThrAlaHisValHisGlyAlaIleTyrLysGlnArgG

4321 GGTGCTTACCTCAGCAGGGAGGGAAATAAAGAACAAAGAGGAAATTCTAAGCCTATTAG 4380
lyLeuLeuThrSerAlaGlyArgGluIleLysAsnLysGluGluIleLeuSerLeuLeuG

4381 AAGCCTTACATTTGCCAAAAGGCTAGCTATTATACACTGTCCTGGACATCAGAAAGCCA 4440
luAlaLeuHisLeuProLysArgLeuAlaIleIleHisCysProGlyHisGlnLysAlaL

4441 AAGATCTCATATCTAGAGGGAACCAGATGGCTGACCGGGTTGCCAAGCAGGCAGCCCAGG 4500
ysAspLeuIleSerArgGlyAsnGlnMetAlaAspArgValAlaLysGlnAlaAlaGlnA

4501 CTGTAAACCTTCTGCCTATAATAGAAACGCCCAAGCCCCAGAACCCAGACGACAGTACA 4560
laValAsnLeuLeuProIleIleGluThrProLysAlaProGluProArgArgGlnTyrT

4561 CCCTAGAAGACTGGCAAGAGATAAAAAAGATAGACCAGTTCTCTGAGACTCCGGAGGGGA 4620
hrLeuGluAspTrpGlnGluIleLysLysIleAspGlnPheSerGluThrProGluGlyT

4621 CCTGCTATACCTCATATGGGAAGGAAATCCTGCCCCACAAAGAAGGGTTAGAATATGTCC 4680
hrCysTyrThrSerTyrGlyLysGluIleLeuProHisLysGluGlyLeuGluTyrValG

4681 AACAGATACATCGTCTAACCCACCTAGGAATAAACACCTGCAGCAGTTGGTCAGAACAT 4740
lnGlnIleHisArgLeuThrHisLeuGlyThrLysHisLeuGlnGlnLeuValArgThrS

4741 CCCCTTATCATGTTCTGAGGCTACCAGGAGTGGCTGACTCGGTGGTCAAACATTGTGTGC 4800
erProTyrHisValLeuArgLeuProGlyValAlaAspSerValValLysHisCysValP

4801 CCTGCCAGCTGGTTAATGCTAATCCTTCCAGAATACCTCCAGGAAAGAGACTAAGGGGAA 4860
roCysGlnLeuValAsnAlaAsnProSerArgIleProProGlyLysArgLeuArgGlys

4861 GCCACCCAGGCGCTCACTGGGAAGTGGACTTCACTGAGGTAAAGCCGGCTAAATACGGAA 4920
erHisProGlyAlaHisTrpGluValAspPheThrGluValLysProAlaLysTyrGlyA

Figure 3 cont.

| | | |
|------|--|------|
| 4921 | ACAAATATCTATTGGTTTTTGTAGACACCTTTTCAGGATGGGTAGAGGCTTATCCTACTA snLysTyrLeuLeuValPheValAspThrPheSerGlyTrpValGluAlaTyrProThrL | 4980 |
| 4981 | AGAAAGAGACTTCAACCGTGGTGGCTAAGAAAATACTGGAGGAAATTTTCCAAGATTTG ysLysGluThrSerThrValValAlaLysLysIleLeuGluGluIlePheProArgPheG | 5040 |
| 5041 | GAATACCTAAGGTAATAGGGTCAGACAATGGTCCAGCTTTCGTTGCCAGGTAAAGTCAGG lyIleProLysValIleGlySerAspAsnGlyProAlaPheValAlaGlnValSerGlnG | 5100 |
| 5101 | GACTGGCCAAGATATTGGGGATTGATTGAAAACCTGCATTGTGCATACAGACCCCAAAGCT lyLeuAlaLysIleLeuGlyIleAspTrpLysLeuHisCysAlaTyrArgProGlnSerS | 5160 |
| 5161 | CAGGACAGGTAGAGAGGATGAATAGAACCATTAAGAGACCCTTACCAAATTGACCACAG erGlyGlnValGluArgMetAsnArgThrIleLysGluThrLeuThrLysLeuThrThrG | 5220 |
| 5221 | AGACTGGCATTAAATGATTGGATGGCTCTCCTGCCCTTTGTGCTTTTTAGGGTGAGGAACA luThrGlyIleAsnAspTrpMetAlaLeuLeuProPheValLeuPheArgValArgAsnT | 5280 |
| 5281 | CCCCTGGACAGTTTGGGCTGACCCCTATGAATTGCTCTACGGGGGACCCCCCGTTGG hrProGlyGlnPheGlyLeuThrProTyrGluLeuLeuTyrGlyGlyProProProLeuA | 5340 |
| 5341 | CAGAAATTGCCTTTGCACATAGTGCTGATGTGCTGCTTTCCAGCCTTTGTTCTCTAGGC laGluIleAlaPheAlaHisSerAlaAspValLeuLeuSerGlnProLeuPheSerArgL | 5400 |
| 5401 | TCAAGGCGCTCGAGTGGGTGAGGCAGCGAGCGTGAAGCAGCTCCGGGAGGCCTACTCAG euLysAlaLeuGluTrpValArgGlnArgAlaTrpLysGlnLeuArgGluAlaTyrSerG | 5460 |
| 5461 | GAGGAGACTTGCAAGTTCACATCGCTTCCAAGTTGGAGATTCAGTCTATGTTAGACGCC lyGlyAspLeuGlnValProHisArgPheGlnValGlyAspSerValTyrValArgArgH | 5520 |
| 5521 | ACCGTGCAGGAAACCTCGAGACTCGGTGGAAGGGACCTTATCTCGTACTTTTGACCACAC lsArgAlaGlyAsnLeuGluThrArgTrpLysGlyProTyrLeuValLeuLeuThrThrP | 5580 |
| 5581 | CAACGGCTGTGAAAGTCGAAGGAATCCCCACCTGGATCCATGCATCCCACGTTAAGCCGG roThrAlaValLysValGluGlyIleProThrTrpIleHisAlaSerHisValLysProA MetHisProThrLeuSerArg | 5640 |
| 5641 | CGCCACCTCCCGACTCGGGGTGGAGAGCCGAAAAGACTGAGAATCCCCTTAAGCTTCGCC laProProProAspSerGlyTrpArgAlaGluLysThrGluAsnProLeuLysLeuArgL ArgHisLeuProThrArgGlyGlyGluProLysArgLeuArgIleProLeuSerPheAla | 5700 |
| 5701 | TCCATCGCCTGGTTCCTTACTCTAACAATAACTCCCCAGGCCAGTAGTAAACGCCTTATA euHisArgLeuValProTyrSerAsnAsnAsnSerProGlyGlnEnd SerIleAlaTrpPheLeuThrLeuThrIleThrProGlnAlaSerSerLysArgLeuIle | 5760 |
| 5761 | GACAGCTCGAACCCCATAGACCTTTATCCCTTACCTGGCTGATTATTGACCCTGATACG AspSerSerAsnProHisArgProLeuSerLeuThrTrpLeuIleIleAspProAspThr | 5820 |

17/22

Figure 3 cont

5821 GGTGTCACCTGTAAATAGCACTCGAGGTGTTGCTCCTAGAGGCACCTGGTGGCCTGAACTG 5880
GlyValThrValAsnSerThrArgGlyValAlaProArgGlyThrTrpTrpProGluLeu

5881 CATTCTCTGCCTCCGATTGATTAACCCCGCTGTTAAAAGCACACCTCCCAACCTAGTCCGT 5940
HisPheCysLeuArgLeuIleAsnProAlaValLysSerThrProProAsnLeuValArg

5941 AGTTATGGGTTCTATTGCTGCCCAGGCACAGAGAAAGAGAAATACTGTGGGGGTTCTGGG 6000
SerTyrGlyPheTyrCysCysProGlyThrGluLysGluLysTyrCysGlyGlySerGly

6001 GAATCCTTCTGTAGGAGATGGAGCTGCGTCACCTCCAACGATGGAGACTGGAAATGGCCG 6060
GluSerPheCysArgArgTrpSerCysValThrSerAsnAspGlyAspTrpLysTrpPro

6061 ATCTCTCTCCAGGACCGGTAAAATTCTCCTTTGTCAATTCCGGCCCGGGCAAGTACAAA 6120
IleSerLeuGlnAspArgValLysPheSerPheValAsnSerGlyProGlyLysTyrLys

6121 ATGATGAACTATATAAGATAAGAGCTGCTCCCCATCAGACTTAGATTATCTAAAGATA 6180
MetMetLysLeuTyrLysAspLysSerCysSerProSerAspLeuAspTyrLeuLysIle

6181 AGTTTCACTGAAAGGAAAACAGGAAAATATTCAAAGTGGATAAATGGTATGAGCTGGGG 6240
SerPheThrGluArgLysThrGlyLysTyrSerLysValAspLysTrpTyrGluLeuGly

6241 AATAGTTTTTTATTATATGGCGGGGGAGCAGGGTCCACTTTAACCATTGCGCTTAGGATA 6300
AsnSerPheLeuLeuTyrGlyGlyGlyAlaGlySerThrLeuThrIleArgLeuArgIle

6301 GAGACGGGGACAGAACCCCTGTGGCAATGGGACCCGATAAAGTACTGGCTGAACAGGGG 6360
GluThrGlyThrGluProProValAlaMetGlyProAspLysValLeuAlaGluGlnGly

6361 CCCCCGGCCCTGGAGCCACCGCATAACTTGCCSGTGCCCCAATTAACCTCGCTGCGGCCCT 6420
ProProAlaLeuGluProProHisAsnLeuProValProGlnLeuThrSerLeuArgPro

6421 GACATAACACAGCCGCTAGCAACAGTACCACTGGATTGATTCTACCAACACGCGCTAGA 6480
AspIleThrGlnProProSerAsnSerThrThrGlyLeuIleProThrAsnThrProArg

6481 AACTCCCCAGGTGTTCTGTTAAGACAGGACAGAGACTCTTCAGTCTCATCCAGGGAGCT 6540
AsnSerProGlyValProValLysThrGlyGlnArgLeuPheSerLeuIleGlnGlyAla

6541 TTCCAAGCCATCAACTCCACCGACCCTGATGCCACTTCTTCTTGTGGCTTTGTCTATCC 6600
PheGlnAlaIleAsnSerThrAspProAspAlaThrSerSerCysTrpLeuCysLeuSer

6601 TCAGGGCCTCCTTATTATGAGGGGATGGCTAAAGAAAGAAAATTCAATGTGACCAAAGAG 6660
SerGlyProProTyrTyrGluGlyMetAlaLysGluArgLysPheAsnValThrLysGlu

6661 CATAGAAATCAATGTACATGGGGGTCCCGAAATAAGCTTACCCTCACTGAAGTTTCCGGG 6720
HisArgAsnGlnCysThrTrpGlySerArgAsnLysLeuThrLeuThrGluValSerGly

6721 AAGGGGACATGCATAGGAAAAGCTCCCCCATCCCAACACCTTTGCTATAGTACTGTG 6780
LysGlyThrCysIleGlyLysAlaProProSerHisGlnHisLeuCysTyrSerThrVal

Figure 3 cont.

18/22

| | | |
|------|--|------|
| 6781 | GTTTATGAGCAGGCCTCAGAAAATCAGTATTTAGTACCTGGTTATAACAGGTGGTGGCA ValTyrGluGlnAlaSerGluAsnGlnTyrLeuValProGlyTyrAsnArgTrpTrpAla | 6840 |
| 6841 | TGCAATACTGGGTAAACCCCTGTGTTTCCACCTCAGTCTTCAACCAATCCAAAGATTTT CysAsnThrGlyLeuThrProCysValSerThrSerValPheAsnGlnSerLysAspPhe | 6900 |
| 6901 | TGTGTCATGGTCCAAATCGTCCCCGAGTGTACTACCATCCTGAGGAAGTGGTCCTTGAT CysValMetValGlnIleValProArgValTyrTyrHisProGluGluValValLeuAsp | 6960 |
| 6961 | GAATATGACTATCGGTATAACCGACCAAAAAGAGAACCCGTATCCCTTACCCTAGCTGTA GluTyrAspTyrArgTyrAsnArgProLysArgGluProValSerLeuThrLeuAlaVal | 7020 |
| 7021 | ATGCTCGGATTAGGGACGGCCGTTGGCGTAGGAACAGGGACAGCTGCCCTGATCACAGGA MetLeuGlyLeuGlyThrAlaValGlyValGlyThrGlyThrAlaAlaLeuIleThrGly | 7080 |
| 7081 | CCACAGCAGCTAGAGAAAGGACTTGGTGAGCTACATGCGGCCATGACAGAAGATCTCCGA ProGlnGlnLeuGluLysGlyLeuGlyGluLeuHisAlaAlaMetThrGluAspLeuArg | 7140 |
| 7141 | GCCTTAAAGGAGTCTGTTAGCAACCTAGAAGAGTCCCTGACTTCCTTGTCTGAAGTGGTT AlaLeuLysGluSerValSerAsnLeuGluGluSerLeuThrSerLeuSerGluValVal | 7200 |
| 7201 | CTACAGAACCGAGGGGATTAGATCTGCTGTTTCTAAGAGAAGGTGGGTTATGTGCAGCC LeuGlnAsnArgArgGlyLeuAspLeuLeuPheLeuArgGluGlyGlyLeuCysAlaAla | 7260 |
| 7261 | TTAAAAGAAGATGTTGCTTCTATGTAGATCACTCAGGAGCCATCAGAGACTCCATGAAC LeuLysGluGluCysCysPheTyrValAspHisSerGlyAlaIleArgAspSerMetAsn | 7320 |
| 7321 | AAGCTTAGAAAAAGTTAGAGAGGCGTCGAAGGGAAAGAGAGGCTGACCAGGGGTGGTTT LysLeuArgLysLysLeuGluArgArgArgArgGluArgGluAlaAspGlnGlyTrpPhe | 7380 |
| 7381 | GAAGGATGGTTCAACAGGTCTCCTTGGATGACCACCCTGCTTTCTGCTCTGACGGGGCCC GluGlyTrpPheAsnArgSerProTrpMetThrThrLeuLeuSerAlaLeuThrGlyPro | 7440 |
| 7441 | CTAGTAGTCCTGCTCCTGTTACTTACAGTTGGGCCTTGCTTAATTAATAGGTTTGTGCC LeuValValLeuLeuLeuLeuLeuThrValGlyProCysLeuIleAsnArgPheValAla | 7500 |
| 7501 | TTTGTTAGAGAACGAGTGAGTGCAGTCCAGATCATGGTACTTAGGCAACAGTACCAAGGC PheValArgGluArgValSerAlaValGlnIleMetValLeuArgGlnGlnTyrGlnGly | 7560 |
| 7561 | CTTCTGAGCCAAGGAGAACTGACCTCTAGCCTTCCCAGTTCTAAGATTAGAACTATTAA LeuLeuSerGlnGlyGluThrAspLeuEnd | 7620 |
| 7621 | CAAGACAAGAAGTGGGGAATGAAAGGATGAAAATGCAACCTAACCCCTCCCAGAACCAGG LeuLeuSerGlnGlyGluThrAspLeuEnd | 7680 |
| 7681 | AAGTTAATAAAAAGCTCTAAATGCCCCCGAATTACAGACCCTGCTGGCTGCCAGTAAATA LeuLeuSerGlnGlyGluThrAspLeuEnd | 7740 |

Figure 3 cont.

19/22

7741 GGTAGAAGGTCACACTTCCTATTGTTCCAGGGCCTGCTATCCTGGCCTAAGTAAGATAAC 7800
7801 AGGAAATGAGTTGACTAATCGCTTATCTGGATTCTGTAAACTGACTGGCACCATAGAAG 7860
7861 AATTGATTACACATTGACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCC 7920
7921 AGGAGCCCACGCAGATGCGGACCTCCGGAGCTATTTTAAATGATTGGTCCACGGAGCGC 7980
7981 GGGCTCTCGATATTTTAAATGATTGGTCCATGGAGCGCGGGCTCTCGATATTTTAAAT 8040
8041 GATTGGTTTGTGACGCACAGGCTTTGTTGTGAACCCCATAAAAGCTGTCCCGATTCCGCA 8100
8101 CTCGGGGCCGAGTCCTCTACCCCTGCGTGGTGTACGACTGTGGGCCCCAGCGCGCTTGG 8160
8161 AATAAAATCCTCTTGCTGTTTGCATCAAAAAAAAAAAAAAAAAAAAAA 8209

20/22

Figure 4.

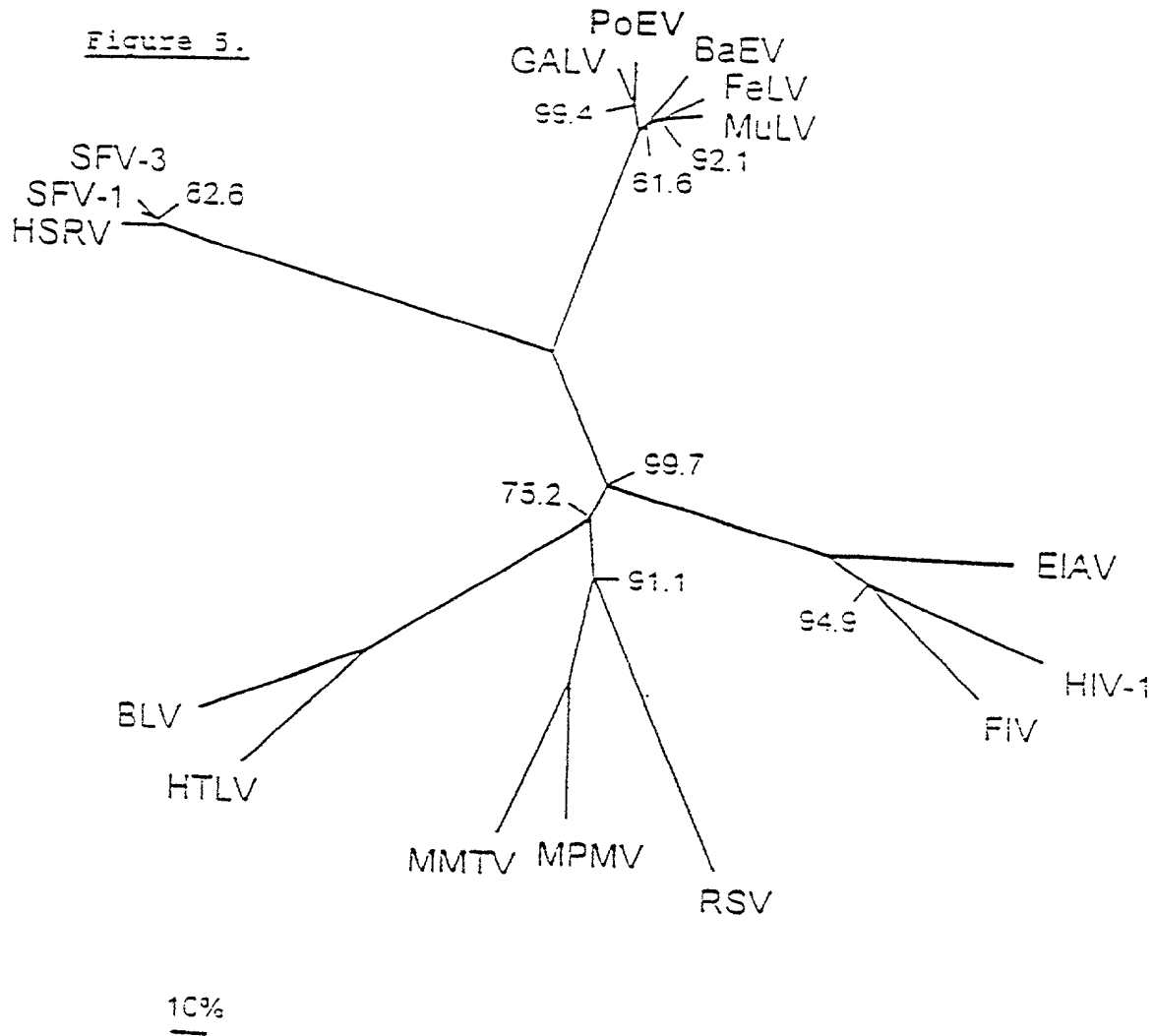
The same nucleotide sequence as represented by bases 5260 to 8210 in Figure 3 is also representative for this Figure, with the following changes:

| <u>Position</u> | <u>Change</u> |
|-----------------|---------------|
| 5273 | G-T |
| 5341 | C-T |
| 5351 | C-T |
| 5353 | T-C |
| 5356 | C-T |
| 5426 | G-A |
| 5464 | Insertion AGA |
| 5607 | C-T |
| 5638 | C-T |
| 5792 | T-C |
| 6191 | Insertion AA |
| 6253 | T-A |
| 6255 | Insertion A |
| 6900 | C-G |

Such nucleotide changes result in the following amino acid changes in the ENV polypeptide.

| <u>Position</u> | <u>Change</u> |
|-----------------|-------------------|
| 7 | R-W |
| 192 | R-K |
| 193 | Deletion |
| 194 | Deletion |
| 197 | Y-Q |
| 198 | S-E |
| 199 | K-N |
| 200 | V-I |
| 201 | D-Q |
| 204 | Y-I |
| 205 | E-N |
| 206 | Insertions: G,M,S |
| 206 | L-W |
| 208 | N-I |
| 209 | S-V |
| 211 | L-Y |
| 212 | L-K |
| 427 | F-L |

21/22

Figure 5.

MuLV murine leukaemia virus
 FeLV feline leukaemia virus
 GaLV gibbon ape leukaemia virus
 SFV-1 simian foamy virus 1
 SFV-3 simian foamy virus 3
 HSRV human foamy virus
 BLV Bovine leukaemia virus
 HTLV human T-cell leukaemia virus
 MMTV murine mammary tumour virus
 MPMV Mason Pfizer monkey virus
 RSV Rous sarcoma virus
 FIV feline immunodeficiency virus
 HIV human immunodeficiency virus
 EIAV equine infectious anaemia virus

SUBSTITUTE SHEET (RULE 26)

PPT | U3 OCT-1 c-Myb LyF-1 E47

1 AAGAAGTGGGGAATGAAAGGATGAAAATGCAACCTAACCCCTCCCAGAACC

ETS-1 AP-4

51 CAGGAAGTTAATAAAAAAGCTCTAAATGCCCCCGAATTMCAGACCCTGCTG

NF-1 AP-1/TR

101 GCTGCCAGTAAATAGGTAGTAAGGTCACACTTCCTATTGTTCCAGGGCCTG

ETS-1/GATA GATA ETS-1 c-Myb AP-1 GATA

151 CTATCCTGGCCTAAGTAAGATAACAGGAAATGAGTTGACTAATCGCTTAT

E47 AP-1

201 CTGGATTCTGTAAAACACTGACTGGCACCATAGAAGAATTGATTACACATTG

AP-1 AP-1/GATA c-Myb AP-1

251 ACAGCCCTAGTGACCTATCTCAACTGCAATCTGTCACTCTGCCCAGGAGC

E47 ETS-1 → CCAAT

301 CCACGCAGATGCGGACCTCCGGAGCTATTTTAAATGATTGGTCCACGGA

GATA → CCAAT ←

351 GCGCGGGCTCTCGATATTTTAAATGATTGGTCCATGGAGCGCGGGCTCT

GATA CCAAT ← AP-1/CREB

401 CGATATTTTAAATGATTGGTTTGTGACGCACAGGCTTTGTTGTGAACCC

TATA U3 | R

451 CATAAAAGCTGTCCCGATTCCGCACTCGGGGCCGAGTCCTCTACCCCTG

PADS polyA

501 CGTGGTGTACGACTGTGGGCCCCAGCGCGCTTGGAATAAAAAATCCTCTTG

R | U5

551 CTGTTTGCATCAAGACCGCTTCTYGTGAGTGATTGTTGGGGTGTGCCTCTT

U5 | PBS

601 CCGAKCCCGGACGAGGGGGATTGTTCTTTTACTGGCCTTTCATTGGTGC

651 GTTGGCCGGGAAATCCTGCGACC

DECLARATION FOR PATENT APPLICATION

As a below-named inventor, I hereby declare that:

My residence, post office address, and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and sole inventor (if more than one name is listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

"PORCINE RETROVIRUS"

the specification of which (check one): ☐ is attached hereto
☐ was filed on _____ as Serial No. _____, and
was amended on _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with 37 CFR § 1.56(a).

Prior Foreign Application(s): I hereby claim foreign priority benefits under 35 USC § 119 of any foreign application(s) for patent or inventor's certificate listed below, and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

| Application No. | Country | Day/Month/Year Filed | Priority Claimed |
|-----------------|---------|----------------------|---|
| GB9608164.1 | GB | 19 APRIL 1996 | <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No |
| GB9702668.6 | GB | 10 FEBRUARY 97 | <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No |
| | | | <input type="checkbox"/> Yes <input type="checkbox"/> No |

I hereby appoint:

| | | | |
|--------------------|------------------------|----------------------|-------------------------|
| Donald W. Muirhead | Reg. No. <u>33,978</u> | John C. Gorecki | Reg. No. <u>38,471</u> |
| Beth E. Arnold | Reg. No. <u>35,430</u> | Edward J. Kelly | Reg. No. <u>38,936</u> |
| Matthew P. Vincent | Reg. No. <u>36,709</u> | Sayoko Blodgett-Ford | Reg. No. <u>P-40516</u> |
| Charles H. Cella | Reg. No. <u>38,099</u> | | |

DECLARATION FOR PATENT APPLICATION

of the Patent Group of Foley, Hoag & Eliot LLP my attorneys and/or agents to prosecute this application and transact all business in the U.S. Patent and Trademark Office connected therewith.

Send Correspondence to: Patent Group :
Foley, Hoag & Eliot LLP
One Post Office Square
Boston, MA 02109

Direct Telephone Calls to:

Matthew P. Vincent, Ph.D. (617) 832-1299

I hereby claim the benefit under 35 USC §120 of any United States application(s) or any PCT international application(s) designating the United States of America listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application in the manner provided by 35 USC §112, first paragraph, I acknowledge the duty to disclose material information as defined in 37 CFR §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

| U.S. Application Serial No. | U.S. Filing Date | Status (patented, pending, abandoned) |
|-----------------------------|------------------|---------------------------------------|
| | | |
| | | |

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements are made with the knowledge that wilful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 USC §1001, and that such wilful false statements may jeopardize the validity of the application or any patent issued thereon.

100

| | |
|--|--|
| Full name of sole or first inventor: | <u>DANIEL NORMAN GALBRAITH</u> |
| Inventor's signature: | <i>Daniel Galbraith</i> |
| Residence Address: | 26 DON STREET, GLASGOW G33 2DH, UNITED KINGDOM <u>GBX</u> |
| Citizenship: | BRITISH |
| Post Office Address (if different from above): | |

200

| | |
|--|--|
| Full name of second inventor: | <u>CHRISTINE HAWORTH</u> |
| Inventor's signature: | <i>Christine Haworth</i> |
| Residence Address: | 139 MARLBOROUGH AVENUE, BROOMHILL, GLASGOW G11 7JE, UNITED KINGDOM <u>GBX</u> |
| Citizenship: | BRITISH |
| Post Office Address (if different from above): | |

300

[Signature]

GBX

BRITISH

Post Office Address (if different from above):

400

10/10/10

CBX

BRITISH

Post Office Address (if different from above):

[illegible]